GenCore version 5_1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 16, 2003, 08:59:57 ; Search time 26 Seconds
(without alignments)
520.049 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLRLFSTNHQSLLLPSSL.....KAVMVEEEPKKRGFFSFFGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	799.5	48.3	286	1	MIND_MESVI	Q9mum5 mesostigma
2	792	47.9	282	1	MIND_CHLVU	P56346 chlorella v
3	694.5	42.0	269	1	MIND_GUITH	O78436 guillardia
4	643	38.9	266	1	MIND_SYNY3	Q55900 synechocyst
5	523	31.6	268	1	MIND_BACSU	Q01464 bacillus su
6	506.5	30.6	268	1	MIND_HELPY	O25098 helicobacte
7	500.5	30.3	268	1	MIND_HELPJ	Q9zma8 helicobacte
8	479	29.0	269	1	MIND_ECOLI	P18197 escherichia
9	475.5	28.7	271	1	MIND_THEMA	Q9x2i3 thermotoga
10	460	27.8	269	1	MIND_BUCAI	P57411 buchnera ap
11	300	18.1	264	1	Y547_METJA	Q57967 methanococc
12	242	14.6	263	1	Y169_METJA	Q57633 methanococc
13	214	12.9	298	1	YLXH_BACSU	P40742 bacillus su
14	213	12.9	353	1	MRP_SYNY3	P53383 synechocyst
15	211	12.8	364	1	MRP_AQUAE	O66946 aquifex aeo
16	209	12.6	368	1	MRP_HELPJ	Q9zmm5 helicobacte
17	206	12.5	368	1	MRP_HELPY	O24999 helicobacte
18	197.5	11.9	350	1	MRP_DEIRA	Q9rvm9 deinococcus
19	196.5	11.9	295	1	YH95_PYRAB	Q9v0d9 pyrococcus
20	193.5	11.7	295	1	Y949_PYRHO	O58667 pyrococcus
21	191	11.5	290	1	Y283_METJA	Q57731 methanococc
22	190	11.5	381	1	MRP_MYCTU	O33225 mycobacteri
23	183.5	11.1	253	1	Y685_METJA	Q58098 methanococc
24	183	11.1	253	1	SOJ_BACSU	P37522 bacillus su
25	181	10.9	287	1	YCAB_PSEFR	P72190 pseudomonas
26	177.5	10.7	383	1	MRP_MYCLE	P53382 mycobacteri
27	173.5	10.5	271	1	NUB2_HUMAN	Q9y5y2 homo sapien
28	173.5	10.5	369	1	MRP_ECOLI	P21590 escherichia
29	165	10.0	257	1	Y823_METJA	Q58233 methanococc
30	164.5	9.9	267	1	PARA_CAUCR	O05189 caulobacter
31	164.5	9.9	275	1	NUB2_MOUSE	Q9r061 mus musculu
32	161	9.7	263	1	Y924_METJA	Q58334 methanococc
33	160.5	9.7	287	1	BCHL_HELMO	Q9zgf0 heliobacill
34	153.5	9.3	290	1	BCHL_RHORU	Q9l8j5 rhodospiril
35	152	9.2	352	1	MRP_BACSU	P50863 bacillus su
36	150.5	9.1	253	1	SOJ_BACHD	Q9k5n0 bacillus ha
37	149.5	9.0	263	1	YGI1_PSEPU	P31856 pseudomonas
38	146.5	8.9	276	1	BCHL_CHLTE	Q9f714 chlorobium
39	146.5	8.9	293	1	YIA3_YEAST	P40558 saccharomyc
40	144.5	8.7	370	1	MRP_HAEIN	P45135 haemophilus
41	143	8.6	320	1	NUB1_HUMAN	P53384 homo sapien
42	140.5	8.5	296	1	NIFH_ALCFA	Q44044 alcaligenes
43	140.5	8.5	320	1	NUB1_MOUSE	Q9r060 mus musculu
44	140	8.5	253	1	SOJ_TREPA	O83296 treponema p
45	137	8.3	287	1	NIFH_FRAAL	P08925 frankia aln

ALIGNMENTS

RESULT 1
MIND_MESVI
ID MIND_MESVI STANDARD; PRT; 286 AA.
AC Q9MUM5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative septum site-determining protein minD.
GN MIND.
OS Mesostigma viride.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC Mesostigmatales; Mesostigmataceae; Mesostigma.
OX NCBI_TaxID=41882;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIES-296;
RX MEDLINE=20150907; PubMed=10688199;
RA Lemieux C., Otis C., Turmel M.;
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT branch of green plant evolution.";
RL Nature 403:649-652(2000).
CC -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC DIVISION SITE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC ----------

CC ----------
DR EMBL; AF166114; AAF43873.1; -.
DR InterPro; IPR000707; ATPase_ParA.
DR Pfam; PF00991; ParA; 1.
KW Cell division; Septation; ATP-binding; Chloroplast.
FT NP_BIND 25 32 ATP (POTENTIAL).
SQ SEQUENCE 286 AA; 31627 MW, 25D9B8FB2258E3F9 CRC64;

Query Match 48.3%; Score 799.5; DB 1; Length 286;
Best Local Similarity 60.4%; Pred. No. 8.4e-57;
Matches 166; Conservative 39; Mismatches 61; Indels 9; Gaps 4;

```
Qy      46 QFNRKPELAGET-PRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDL 104
           | |:  |    |  | :|||||||||||||||||:|:|:|| |: |  ||||:|||||||
Db       4 QINKDGEKKNSTDTRTIVITSGKGGVGKTTTTANLGMSIARLGYKVALIDADVGLRNLDL 63

Qy     105 LLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLV 164
           |||||||| || :|| |:| ||||:|||||| || ||| | :  :    : :| ||
Db      64 LLGLENRVIYTAMEVFEGECCLDQALIRDKRWSNLALLAISKTRQRYHL--TRRNMEMLV 121

Qy     165 DALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECD 224
           |::: |   : :||:|||||||| ||. |: || |||:||||:||::|||||| ||||
Db     122 DSIRLR---NYNFILIDCPAGIDVGFVNAVAPAEEAVVVTTPEITSIRDADRVAGLLEAS 178
```

```
Qy      225 GIPDIKMIVNPVRTPMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNK 284
            || · |:·||||| ||·  |: || ||||||| ·||·||||· || ||||| ||||||
Db      179 GIYEVKLLVNPVRPDMIQKNDMLSVRDVQEMLGIPLLGAIPEDTNVIVSTNRGQPLVLNK 238

Qy      285 PPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRG 319
              ||:|::|| || |||     | :|  |   :|
Db      239 KLTLSGISFENAARRLV---GRKEYLVNLETGNKG 270


RESULT 2
MIND_CHLVU
ID   MIND_CHLVU     STANDARD;      PRT;   282 AA.
AC   P56346;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative septum site-determining protein minD.
GN   MIND.
OS   Chlorella vulgaris.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae, Chlorellales,
OC   Chlorellaceae; Chlorella.
OX   NCBI_TaxID=3077;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN-IAM C-27 / Tamiya;
RX   MEDLINE=97303241; PubMed=9159184;
RA   Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA   Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA   Inamura A., Yoshinaga K., Sugiura M.;
RT   "Complete nucleotide sequence of the chloroplast genome from the
RT   green alga Chlorella vulgaris: the existence of genes possibly
RT   involved in chloroplast division.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB001684; BAA57951.1; -.
DR   InterPro; IPR000707; ATPase_ParA
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Chloroplast
FT   NP_BIND      24     31       ATP (POTENTIAL).
SQ   SEQUENCE   282 AA;  31013 MW;  94010DD45AE4AEC7 CRC64;

  Query Match             47.9%;  Score 792;  DB 1;  Length 282;
  Best Local Similarity   60.4%;  Pred. No. 3.3e-56;
  Matches  165;  Conservative   39;  Mismatches   55;  Indels   14;  Gaps    5;

Qy       59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
            |:|:|||||||||||||||||:|:|:|| |: || :|||||||||||||||||||| ||  ::
Db       17 RVIVITSGKGGVGKTTTTANLGMSIARLGYRVALIDADIGLRNLDLLLGLENRVLYTAMD 76

Qy      119 VINGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTPPEGSPDFI 178
             :· | |||||||·||||| |  || ||| |     |   ::  |:|::|    |    :
Db       77 IVEGQCRLDQALIRDKRWKNLALLAISKNRQK--YNVTRKNMQNLIDSVK---ELGFQFV 131

Qy      179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRT 238
            :||||||| ||| ||  | |||:||||:|||||||| |||| :|| ::|::|||||
Db      132 LIDCPAGIDVGFINAIASAQEAVIVTTPEITAIRDADRVAGLLEANGIYNVKLLVNRVRP 191

Qy      239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
            |||:  ||||| |||||||: ||| ||||: || |||:| ||||||   ||:|:||| ||
Db      192 DMIQKNDMMSVRDVQEMLGIPLLGAIPEDTSVIISTNKGEPLVLNKKLTLSGIAFENAAR 251
```

```
Qy      299 RLV-EQDSMKAVMVEEEPKKRGFF---SFFGG 326
             ||  ||       ::    ··| |     || |
Db      252 RLIGKQD----YFIDLTSPQKGMFQKLQEFFLG 280


RESULT 3
MIND_GUITH
ID   MIND_GUITH     STANDARD;      PRT;   269 AA.
AC   O78436;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Putative septum site-determining protein minD.
GN   MIND.
OS   Guillardia theta (Cryptomonas phi)
OG   Chloroplast.
OC   Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
OX   NCBI_TaxID=55529;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99128221; PubMed=9929392;
RA   Douglas S.E., Penny S.L.;
RT   "The plastid genome of the cryptophyte alga, Guillardia theta:
RT   complete sequence and conserved synteny groups confirm its common
RT   ancestry with red algae.";
RL   J. Mol. Evol. 48:236-244(1999).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF041468; AAC35621.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Chloroplast.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   269 AA;  29455 MW;  BC3C3E954E689EA2 CRC64;

  Query Match             42.0%;  Score 694.5;  DB 1;  Length 269;
  Best Local Similarity   55.9%;  Pred. No. 1.9e-48;
  Matches  146;  Conservative   44;  Mismatches   62;  Indels    9;  Gaps    3;

Qy       59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
            ||||||||||||||| |||·|·:·||: |·     ||||·|||||||||||||| || :|
Db        3 RIVVITSGKGGVGKTTVTANLGMALAQLGYRTALIDADIGLRNLDLLLGLENRVIYTALE 62

Qy      119 VINGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTPPEGSPDFI 178
            |:·|·|||·|||·:||  |  ||  ··  |:|          · ···||:|         |·:
Db       63 VLSGECRLEQALIKDKRQPNLVLLPAAQNRNK--DSVTEEQMKFLVNLL----VNDYDYL 116

Qy      179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRT 238
            :||||||: || ||  ||·||||·| |·||||| |||| |||· |  ·||·|·|||:|
Db      117 LIDCPAGIETGFHNAIGPAQEAIVVTTPEIAAVRDADRVIGLLEANGIKQIKLLVNRLRP 176

Qy      239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
            |:| |||| ||:|  · |·|||| |||||||  ||||||  |  :|||||| |
Db      177 QMVKANDMMSVADVREILAIPLIGVIPEDECVIVSTNRGEPLVLEKNLSLPGLAFEHTAC 236

Qy      299 PLVEQDSMKAVMVEEEPKKRG 319
             ||  |  :   : : :   ||
Db      237 PL---DGQEIEFLDLQSYSRG 254


RESULT 4
MIND_SYNY3
```

```
ID   MIND_SYNY3     STANDARD;      PRT;   266 AA.
AC   Q55900;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR SLL0289.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D64005; BAA10662.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   266 AA;  29047 MW;  4BC37880F086C9BF CRC64;

  Query Match             38.9%;  Score 643;  DB 1;  Length 266;
  Best Local Similarity   52.8%;  Pred. No. 2.4e-44;
  Matches  130;  Conservative   51;  Mismatches   59;  Indels    6;  Gaps    2;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           ||:|:|||||||||||||:| :||| |  || |||| ||||||||||| |: || ::
Db       3 RIIVVTSGKGGVGKTTTTANLGAALARLGKKVVLIDADFGLRNLDLLLGLEQRIVYTAID 62

Qy     119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
           |:  :| :|:|||:||| | ||  :: ||| : :: ||: || :    |:|
Db      63 VLADECTIDKALVKDKRLPNLVLLPAAQNRSK--DAINAEQMQSLVEQLKDK----FDYI 116

Qy     179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIPDIKMIVNRVPT 238
           ||||||||:|||  |: || ||::||||:::|:|||||| |||| : |  | :||||:|
Db     117 IIDCPAGIEAGFRNAVAPAQEAIIVTTPEMSAVRDADRVIGLLEAEDIGKISLIVNRLRP 176

Qy     239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
           :|··  |:|| |: ::| : |:|::|:| ::| ||||:| |||: :  :: ||||.  |
Db     177 EMVQLNQMISVEDILDLLAVPLIGILPDDQKIIISTNKGEPLVMEEKLSVPGLAFQNIAR 236

Qy     299 RLVEQD 304
           ||  ||
Db     237 RLEGQD 242


RESULT 5
MIND_BACSU
ID   MIND_BACSU     STANDARD;      PRT;   268 AA.
AC   Q01464;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR DIVIVB.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=93015732; PubMed=1400225;
RA   Varley A.W., Stewart G.C.;
RT   "The divIVB region of the Bacillus subtilis chromosome encodes
RT   homologs of Escherichia coli septum placement (minCD) and cell shape
RT   (mreBCD) determinants.";
RL   J. Bacteriol. 174:6729-6742(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=93211302; PubMed=8459776;
RA   Lee S., Price C.W.;
RT   "The minCD locus of Bacillus subtilis lacks the minE determinant that
RT   provides topological specificity to cell division.";
RL   Mol. Microbiol. 7:601-610(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93015731; PubMed=1400224;
RA   Levin P.A., Margolis P.S., Setlow P., Losick R., Sun D.;
RT   "Identification of Bacillus subtilis genes for septum placement and
RT   shape determination.";
RL   J. Bacteriol. 174:6717-6728(1992).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M95582; AAA22609.1; -.
DR   EMBL; Z15113; CAA78818.1; -.
DR   EMBL; M96343; AAA22401.1; -.
DR   EMBL; Z99118; CAB14759.1; -.
DR   PIR; S31205; S31205.
DR   PIR; G45239; G45239.
DR   PIR; F45240; F45240.
DR   SubtiList; BG10330; minD.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   268 AA;  29407 MW;  6665E9F693F58A9B CRC64;

  Query Match             31.6%;  Score 523;  DB 1;  Length 268;
  Best Local Similarity   41.1%;  Pred. No. 9.4e-35;
  Matches  111;  Conservative   68;  Mismatches   77;  Indels   14;  Gaps    7;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           :|||||||||||||:||:| :||  |  |  :|  |:||||||:::|||||: |  |:|:
Db       5 IVITSGKGGVGKTTTSANLGTALAILGKRVCLVDTDIGLRNLDVVMGLENRIIYDLVDVV 64
```

```
Qy      121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
             | |:: ||||:|||    :                  ·  ·   |   :|        |.:||
Db       65 EGRCKMHQALVKDKRFDDL-LYLMPAAQTSDKTAVAPEQIKNMVQELKQ----EFDYVII 119

Qy      181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC-DGIRDIKMIVNRVRTD 239
            |||||. |:  |.: |: |::||||:| | |||||  ||||     :         |||.:
Db      120 DCPAGIEQGYKNAVSGADKAIVVTTPEISAVRDADRIIGLLEQEENIEPPRLVVNRIRNH 179

Qy      240 MIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWR 299
             ::|  | |            |     |||  : :| |||::::| | |     :    |     | :|:    | |
Db      180 LMKNGDTMDIDEIVQHLSIDLLGIVADDDEVIKASNHGEPIAMD PKNRASIAYRNIARR 238

Qy      300 LVEQDSMKAVMVEEEPKKRGFF----SFFG 325
              .|       ·:||· |   |         |:|:
Db      239 IL-GESVPLQVLEEQNK--GMMAKIKSFFG 265
```

RESULT 6

```
MIND_HELPY
ID   MIND_HELPY     STANDARD;      PRT;   268 AA.
AC   O25098;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR HP0331.
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=26695 / ATCC 700392;
RX   MEDLINE=97394467; PubMed=9252185;
RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA   Venter J.C.;
RT   "The complete genome sequence of the gastric pathogen Helicobacter
RT   pylori.";
RL   Nature 388:539-547(1997).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE000551; AAD07400.1; -.
DR   TIGR; HP0331; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   268 AA;  29294 MW;  9E14766E8D309476 CRC64;

  Query Match             30.6%;  Score 506.5;  DB 1;  Length 268;
  Best Local Similarity   40.1%;  Pred. No. 1.9e-33;
  Matches  110;  Conservative   65;  Mismatches   80;  Indels   19;  Gaps    5;

Qy       60 IVVITSGKGGVGKTTTTANVGLSLAFYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
              :| |||||||||:|||||: :  ||  |   |||.:| |:||||||.:||||:|  |  |.|
Db        4 VVTITSGKGGVGKSTTTANLAIGLAESGKKVVAVDFDIGLRNLDMILGLENRIVYDVVDV 63

Qy      120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
              .  .| |  |||: ||.  |    |   |:  : |  .       |.  |.:||.      |.|:
Db       64 MEKNCNLSQALITDKKTKNLSFLAASQSKDKNILDKEKVAI--LINALR----ADFDYIL 117

Qy      180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIK------MI 232
             || ||||:·||  ||  |:  |:·| ||:::·|||:|||  |···      |   |      ·|
Db      118 IDSPAGIESGFEHAILHADMALVVVTPEVSSLRDSDRVVGIIDAKSNRAKKGMEVHKHLI 177

Qy      233 VNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLA 292
              :||:· ···  ·|·|· ·|  ··|  | |·|·||||  :|:||:|:|  |::  :    :    |
Db      178 INPLKPELVANGEMISIEEVLKILCLPLIGIIPEDHHIISATNKGEPVI  PTPCESAPA 235

Qy      293 FEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
              :..   |..  ::     |     |    |||||     |
Db      236 YQRITRRILGEE----VEYVEFKAKRGFFSALKG 265
```

RESULT 7

```
MIND_HELPJ
ID   MIND_HELPJ     STANDARD;      PRT;   268 AA.
AC   Q9ZMA8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT 2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR JHP0314.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
OX   NCBI_TaxID=85963;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE001468; AAD05905.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
```

```
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   NP_BIND      10      17       ATP (POTENTIAL).
SQ   SEQUENCE   268 AA;  29265 MW;  1F479C187A201A63 CRC64;

  Query Match             30.3%;  Score 500.5;  DB 1;  Length 268;
  Best Local Similarity   39.8%;  Pred. No. 5.9e-33;
  Matches  109;  Conservative   65;  Mismatches   81;  Indels   19;  Gaps    5;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
           :| |||||||||:|||||: : ||  |  |||:| |:|||||::|||||: |  |:|
Db       4 VVTITSGKGGVGKSTTTANLAIGLAESGKKVVAVDFDIGLRNLDMILGLENRIVYDVVDV 63

Qy     120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
           :  :| | |||: ||:  |   |  |: : |  :      |:  |::||:       |:|:
Db      64 MEKNCNLSQALITDKKTKNLSFLAASQSKDKNILDKEKVAI--LINALR----ADFDYIL 117

Qy     180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEQDGIP-------DIKMI 232
           || ||||::|| ||  |: |:.| ||:...|||:||| |...   |           :|
Db     118 IDSPAGIESGFEHAILHADMALVVVTPEVSSLRDSDRVIGIIDAKSNRAKSGEEVHKHLI 177

Qy     233 VNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLA 292
           :||:: :::   :|:|: :| ::| | |:|:|||| :| :||:| |::  :   :  |
Db     178 INRLKPELVANGEMISIEEVLKILCLPLIGIIPEDHHIISATNKGEPVI--PTDCESAKA 235

Qy     293 FEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
           :::  |:: ::   |  |  ||||||  |
Db     236 YQRITRRILGEE----VEYVEFKAKRGFFSALKG 265


RESULT 8
MIND_ECOLI
ID   MIND_ECOLI     STANDARD;      PRT;   269 AA.
AC   P18197;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR B1175 OR Z1937 OR ECS1669.
OS   Escherichia coli, and
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562, 83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89136010; PubMed=2645057;
RA   de Boer P.A.J., Crossley R.E., Rothfield L.I.;
RT   "A division inhibitor and a topological specificity factor coded for
RT   by the minicell locus determine proper placement of the division
RT   septum in E. coli.";
RL   Cell 56:641-649(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.",
RL   DNA Res. 8:11-22(2001).
RN   [6]
RP   SEQUENCE OF 1-12.
RC   STRAIN=K12 / EMG2;
RX   MEDLINE=97443975; PubMed=9298646;
RA   Link A.J., Robison K., Church G.M.;
RT   "Comparing the predicted and observed properties of proteins encoded
RT   in the genome of Escherichia coli K-12.";
RL   Electrophoresis 18:1259-1313(1997).
RN   [7]
RP   FUNCTION, AND MUTAGENESIS.
RX   MEDLINE=92097557; PubMed=1836760;
RA   de Boer P.A.J., Crossley R.E., Hand A.R., Rothfield L.I.;
RT   "The MinD protein is a membrane ATPase required for the correct
RT   placement of the Escherichia coli division site.";
RL   EMBO J. 10:4371-4380(1991).
RN   [8]
RP   CHARACTERIZATION.
RX   MEDLINE=99238467; PubMed=10220403;
RA   Raskin D.M., de Boer P.A.;
RT   "Rapid pole-to-pole oscillation of a protein required for directing
RT   division to the middle of Escherichia coli.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:4971-4976(1999).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS.
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ.
CC   -!- SUBCELLULAR LOCATION: Inner membrane-associated.
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; J03153; AAB59062.1; -.
DR   EMBL; AE000216; AAC74259.1; -.
DR   EMBL, D90751; BAA36009.1; -.
DR   EMBL, D90752; BAA36022.1; -.
DR   EMBL; AE005334; AAG56026.1; -.
DR   EMBL; AP002555; BAB35092.1; -.
DR   PIR; B31877; CCECID.
DR   SWISS-2DPAGE; P18197; COLI.
```

```
DR   EcoGene; EG10597; minD.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR003348; ArsA_ATPase.
DR   Pfam; PF00991; ParA; 1.
DR   Pfam; PF02374; ArsA_ATPase; 1.
KW   Cell division; Septation; ATP-binding; Inner membrane;
KW   Complete proteome.
FT   INIT_MET      0      0
FT   NP_BIND       9     16       ATP (POTENTIAL).
FT   MUTAGEN      14     14       G->S: LESS EFFECTIVE THEN WILD-TYPE.
FT   MUTAGEN      15     15       K->Q: LOSS OF ACTIVITY.
FT   MUTAGEN      15     16       KT->QR: LOSS OF ACTIVITY.
SQ   SEQUENCE   269 AA;  29483 MW;  CD1B29A476A190B1 CRC64;

  Query Match             29.0%;  Score 479;  DB 1;  Length 269;
  Best Local Similarity   39.9%;  Pred. No. 3.1e-31;
  Matches  113;  Conservative   58;  Mismatches   82;  Indels   30;  Gaps    8;

Qy         59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
              ||:|:||||||||||| |   ||: |    | :| | |:|||:|: : | || |
Db          2 RIIVVTSGKGGVGKTTSSAAIATGLAQKGKKTVVIDFDIGLRNLDLIMGCERRVVYDFVN 61

Qy        119 VINGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
              || ||  |:||: :||  |   |   |: | |  :   | |   :: :| ||      :||
Db         62 VIQGDATLNQALIKDKRTENLYILPASQTRDKDALTREGVAK--VLDDLKAM---DFEFI 116

Qy        179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC------DGIRDIK-- 230
              : |  ||||  |   |.  :|:||.  | |  :  :|| ||: |:|           :  ||
Db        117 VCDSPAGIETGALMALYFADEAIITTNPEVSSVRDSDRILGILASKSRRAENGEEPIKEH 176

Qy        231 MIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAG 290
              :: :    ||  |: |   :   |:  | |:||||||  |:  :| :| |  :|    ||
Db        177 LLLTRYNPGRVSRGDMLSMEDVLEILRIKLVGVIPEDQSVLRASNQGEPVILDINAD-AG 235

Qy        291 LAFEQAAWRLVEQDSMKAVMVEEEP   ----KKPGFFS-FFGG 326
              |.        |.  :  :|| |        :|  | |    ||
Db        236 KAY-         ADTVEPLLGEEPPFPFIEEEKKGFLKRLFGG 269

RESULT 9
MIND_THEMA
ID   MIND_THEMA     STANDARD;      PRT;   271 AA.
AC   Q9X2I3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR TM1870.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OC   Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE001824; AAD36932.1; -.
DR   TIGR; TM1870; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   271 AA;  29483 MW;  EC76D3F15D64A4CA CRC64;

  Query Match             28.7%;  Score 475.5;  DB 1;  Length 271;
  Best Local Similarity   40.9%;  Pred. No. 5.9e-31;
  Matches  105;  Conservative   47;  Mismatches   68;  Indels   37;  Gaps    5;

Qy         60 IVVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLPNLDLLLGLENRVNYTCVEV 119
               :|:|||||||||| |||:| :||: |  | |||:|:|||:||||| || :.|
Db          4 VIVVTSGKGGVGKTTITANLGCALAKLGEKVCLIDADIGLKNLDIVLGLENRIVYTMIDV 63

Qy        120 INGDCPLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSP---- 175
               :||  :|:|: |   | | |:  | |                     : |:  ||
Db         64 VNGKVSPQEALVKHKMLKNLYLL----PASQ---------------IATKEMISPNDMK 103

Qy        176 ---------DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDG 225
                       |:|||:|||:|  |: ||  |:   :  |||| .  | |||| |||  |
Db        104 AIVKELIPHFDYIIIDSPAGIERGFRNAVAPAERVLVVTTPELPAISDADRVIGLLENFG 163

Qy        226 IPD--IKMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLN 283
                 |  |: | ||  |:|   |     |    |  |:  ||:  ::| :|| | |: ||
Db        164 FSDEKINVIINRFKPHMVKKGEMLTTDDIKHTLSLEIIAVIPDSEDIIVASNTGIPVSLN 223

Qy        284 KPPTLAGLAFEQAAWRL 300
                  ::  || | |:
Db        224 GNSRIS KNFENLARRI 239

RESULT 10
MIND_BUCAI
ID   MIND_BUCAI     STANDARD;      PRT;   269 AA.
AC   P57411;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Septum site-determining protein minD (Cell division inhibitor minD).
GN   MIND OR BU326.
OS   Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS   symbiotic bacterium).
OC   Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX   NCBI_TaxID=118099;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tokyo 1998;
RX   MEDLINE=20445173; PubMed=10993077;
RA   Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT   "Genome sequence of the endocellular bacterial symbiont of aphids
RT   Buchnera sp. APS.";
RL   Nature 407:81-86(2000).
CC   -!- FUNCTION: ATPASE REQUIRED FOR THE CORRECT PLACEMENT OF THE
CC       DIVISION SITE. CELL DIVISION INHIBITORS MINC AND MIND ACT IN
CC       CONCERT TO FORM AN INHIBITOR CAPABLE OF BLOCKING FORMATION OF THE
CC       POLAR Z RING SEPTUMS. RAPIDLY OSCILLATES BETWEEN THE POLES OF THE
CC       CELL TO DESTABILIZE FTSZ FILAMENTS THAT HAVE FORMED BEFORE THEY
CC       MATURE INTO POLAR Z RINGS (BY SIMILARITY).
CC   -!- SUBUNIT: INTERACTS WITH MINC AND FTSZ (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE PARA FAMILY. MIND SUBFAMILY.
```

```
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP001119; BAB13034.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR003348; ArsA_ATPase.
DR   Pfam; PF00991; ParA; 1.
DR   Pfam; PF02374; ArsA_ATPase; 1.
KW   Cell division; Septation; ATP-binding; Membrane; Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   NP_BIND       9     16       ATP (POTENTIAL)
SQ   SEQUENCE   269 AA;  29805 MW;  4FFD5764549C8B27 CRC64;

  Query Match             27.8%;  Score 460;  DB 1;  Length 269;
  Best Local Similarity   38.3%;  Pred. No. 1e-29;
  Matches  106;  Conservative   60;  Mismatches   91;  Indels   20;  Gaps    5;

Qy      59 PIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           ||:|:|||||||||||::| :|  ||: |   : || |:|||||||::| | || |  :
Db       2 RIIVVTSGKGGVGKTTSSAAIGTGLAQKGKKTIVIDFDIGLRNLDLIMGCERRVVYDFIN 61

Qy     119 VINGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGG--KALEWLVDALKTRPEGSPD 176
           || ||  |:||:::||: :|  :|  |: | |   :   |  | |  |:       : : |
Db      62 VIQGDATLNQAIIKDKKTNNLFILPASQTRDKDALTRIGVEKVLTELI-------KMNFD 114

Qy     177 FIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIK------ 230
           ||| | ||||: | | ||  |:||:: | |:;  ::||:||: |:-      |  |
Db     115 FIICDSPAGIETGAILAIYFADEAIITTNPEVSSVRDSDRILGIISSKSKRAEKNITPIK 174

Qy     231 --MIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNPGFPLVLNKPPTL 288
             :: |     :|  :|:|: || ::| : ::|||||  |:| :|:|  : :|.
Db     175 EYLLLTRYNPPPVKKGEMLSMTDVLDILQIPIIGVIPEDQSVLRASNQGESIILDINSN- 233

Qy     289 AGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFG 325
           || |:     ||: ::  :    || ||     ||
Db     234 AGCAYSDTVNRLLGEE--RHFRFIEEEKKSFLRRLFG 268


RESULT 11
Y547_METJA
ID   Y547_METJA     STANDARD;      PRT;   264 AA.
AC   Q57967;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical ATP-binding protein MJ0547.
GN   MJ0547.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: HIGH, TO M.JANNASCHII MJ0169 AND MJ0410.

CC   -!- SIMILARITY: SOME, TO PROKARYOTIC MIND AND TO THE MRP FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U67504; AAB98539.1; -.
DR   TIGR; MJ0547; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Hypothetical protein; ATP-binding; Complete proteome
FT   NP_BIND      13     21       ATP (POTENTIAL).
SQ   SEQUENCE   264 AA;  27737 MW;  B9AF5108CDD0CEC3 CRC64,

  Query Match             18.1%;  Score 300;  DB 1;  Length 264;
  Best Local Similarity   30.5%;  Pred. No. 6e-17;
  Matches   82;  Conservative   57;  Mismatches  104,  Indels   26,  Gaps    6,

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           : | |||| |||| :||: ::||::|  |  :||: : ||:|::||| :       |
Db       9 IAIASGKGGTGKTTISANLAVALAKFGKKVAVLDADIAMANLELIMGLEGK------PVT 62

Qy     121 NGDCRLDQALVRDKPWSNFELLCISKPPSKLPMG-----FGGKALEWLVDALKTRPEGSP 175
           |    .| .:|  :   |  :       :| |      |      | | : ||   :
Db      63 LNDVLAGKADIKDAIYEGPEGVLV------IPAGVSLEKFRRAKPEKLEEVLKAIHD-LV 115

Qy     176 DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNR 235
           : :||||||||     : ||: |:  ::|  |:|...: || ::  . :  |   | ||||
Db     116 EILIIDCPAGIGKETLIAISSADGLIVVVNPEISSISDALKIIAITKRLGTDIIGAIVNR 175

Qy     236 VRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQ 295
           |  .  .    .  |  .: :| : :.||.|||  | ..   | |||.  | : |  |  :
Db     176 VSNESTE----LGVKAIETILEVPVIGVVPEDPHVRKAAAFGTPLVIMYPDSPAAQAIME 231

Qy     296 AAWRLVEQDSMKAVMVEEEPKKRGFFSFF 324
           | :|:          : : ||  | | |
Db     232 IAAKLIG----AKYEAQLKKKKESFISKF 256


RESULT 12
Y169_METJA
ID   Y169_METJA     STANDARD;      PRT;   263 AA.
AC   Q57633;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical ATP-binding protein MJ0169.
GN   MJ0169.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC   Methanocaldococcaceae; Methanocaldococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96337999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: HIGH, TO M.JANNASCHII MJ0547. SOME, TO MJ0410 AND
CC       TO MJ0924.
```

```
CC   -!- SIMILARITY: SOME, TO PROKARYOTIC MIND AND TO THE MRP FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U67474; AAB98154.1; -.
DR   TIGR; MJ0169; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Hypothetical protein; ATP-binding; Complete proteome.
FT   NP_BIND      10     17       ATP (POTENTIAL).
SQ   SEQUENCE   263 AA;  28950 MW;  292FBE38EE72ACF4 CRC64;

  Query Match            14.6%;  Score 242;  DB 1;  Length 263;
  Best Local Similarity   29.3%;  Pred. No. 2.5e-12;
  Matches   77;  Conservative   59;  Mismatches   99;  Indels   28;  Gaps    9;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
            |. | ||:|||||||.|:.      ||: |  |:|||  |:   ||  :|  :|     ||
Db       5 IITIASGKGGVGKTTTSASLAVALAKLGKKVLAIDGDISMANLGILFNMEKK-KPSLHEV 63

Qy     120 INGDCRLDQALVRD---KRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPD 176
           ::      :| |||    |        |     |  |      |:      : | |   .  |
Db      64 LS-----EEADVRDAIYKHKTGVYVL----PTSLSLEGYKKSDIDLLPDVVNEVADDF-D 113

Qy     177 FIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRV 236
           .:||| ||!:     |    |: :|| ||::  :  || |.     |    |   ·  ::||
Db     114 YVIIDAPAGLNREMATHLAIADKLLLVVTPEMFSIIDAVRLKESAEMAGTPLMGVVLNRV 173

Qy     237 PTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLV---LNKPPTLAGLAF 293
              |  ||  |   :.   ··   |   .|||  |  :  ·   ::    | | :
Db     174 GRDF--GE--MGPDEIEMLIKGKVLVEVPEDENVRSAALKKMSVIEYRKNSPAS---- 223

Qy     294 EQAAWRLVEQDSMKAVMVEEEPK 316
             ||  :|    :   : :|:| |
Db     224 -QAYMKLASIIAGVPIYIEDEIK 245


RESULT 13
YLXH_BACSU
ID   YLXH_BACSU     STANDARD;      PRT;   298 AA.
AC   P40742;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein ylxH.
GN   YLXH.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95078462; PubMed=7987014;
RA   Kirsch M.L., Carpenter P.R., Ordal G.W.;
RT   "A putative ATP-binding protein from the che/fla locus of Bacillus
RT   subtilis.";
RL   DNA Seq. 4:271-275(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA   Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA   Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA   Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the Gram-positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X73682; CAA52034.1; -.
DR   EMBL; Z99112; CAB13514.1; -.
DR   PIR; S34194; S34194.
DR   SubtiList; BG10254; ylxH.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Hypothetical protein; ATP-binding; Complete proteome.
FT   NP_BIND      39     46       ATP (POTENTIAL).
SQ   SEQUENCE   298 AA;  33160 MW;  AF6FE71FA3C1C972 CRC64;

  Query Match            12.9%;  Score 214;  DB 1;  Length 298;
  Best Local Similarity   25.2%;  Pred. No. 5.1e-10;
  Matches   76;  Conservative   56;  Mismatches  119;  Indels   50;  Gaps   10;

Qy      32 VNNPSRRSPIRSVLQFNRK--PELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFS 89
            !     : :|   :       !      ·      · |||||||.  | |. |:|    |
Db       3 MNRYDQAATLRAKMEKRERVLPMVYSQKAKTLAVISGKGGVGKSNITLNMALALQDKGKK 62

Qy      90 VVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPRS 149
            |:  || |.|:  |.|:|:|  |   :  ·:|.      | |.|                 |
Db      63 VLLIDLDIGMGNIDILIG--NSSSATIIDVLTDRKPLLQSL------------------S 102

Qy     150 KLPMGF----GGKALE---------WLVDALKTRPEGSP-DFIIIDCPAGIDAGFITAIT 195
            | |      || |.               |   |    |  |.:.  |  |.     .   |
Db     103 VGPKGLRYISGGTGLDVMFQLDQRKWTFFANELSHALSQFDYVLFDMGAGLSKDQLPFIL 162

Qy     196 PANEAVLVTTPDITALRDA-DRVTGLLECDGIRDIKMIVNRVRTDMIKGEDMMS------ 248
             |   · ::||| |!  ||    |       |       |· ||| |  |  :| |  :
Db     163 SAEDILIITTPEPTAIMDAYSAVKHLVLTENKLSMKVAVNRCR-DQKEGLDAFAPLSRTI 221

Qy     249 --VLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRLVEQDSM 306
              ||||    .   | . :|  |  :     |  :  |  |   :    |  | |.:
Db     222 HMFLDVQ----VQFAGSVSDDVIVSKAVVEQVPFFIKSPQAKASRSVRILADALFEREET 277

Qy     307 K 307
           :
Db     278 R 278


RESULT 14
```

```
MRP_SYNY3
ID   MRP_SYNY3      STANDARD;      PRT;   353 AA.
AC   P53383;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mrp protein homolog.
GN   MRP OR SLR0067.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SIMILARITY: BELONGS TO THE MRP/NBP35 FAMILY OF ATP-BINDING
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D64001; BAA10297.1; -.
DR   InterPro; IPR002744; DUF59.
DR   InterPro; IPR000808; Mrp_family.
DR   Pfam; PF01883; DUF59; 1.
DR   ProDom; PD005595; DUF59; 1.
DR   PROSITE; PS01215; MRP; 1.
KW   ATP-binding; Complete proteome.
FT   NP_BIND     105    112       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  37131 MW;  891DC0E247113841 CRC64;

  Query Match             12.9%;  Score 213;  DB 1;  Length 353;
  Best Local Similarity   26.1%;  Pred. No. 7.7e-10;
  Matches   82;  Conservative   55;  Mismatches  115;  Indels   62;  Gaps   12;

Qy         23 KTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRIVVITSGKGGVGKTTTTANVGLS 82
               :|:  :  | |:|                |.   |: |:|||||||:|     || :
Db         76 KVTAETPQQKSLPDRQS-------------VGQVKNIIAISSGKGGVGKSTVAVNVAVA 121

Qy         83 LARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELL 142
              ||: | :| :|||:   |     :|||        |:|  |      |  :  .   |  :
Db        122 LAQTGAAVGLLDADIYGPNAPTMLGLSG----AAVQVQNS----PQGEVLEPVF-NHGIK 172

Qy        143 CISKPRSKLPMGF----------GGKALEWLVDALKTRPE-GSPDFIIIDCPAGI-DAGF 190
               :|         |||          |   |  ::      :     |: |::|:|  |  |  ||
Db        173 MVS-------MGFLIDPDQPVIWRGPMLNGIIPQFLYQVNWGALDYLIVDMPPGTGDAQL 225

Qy        191 -ITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNR---VRTDMIK---- 242
               :|    |     ||:||||   ||  |         .       ::  |     .   |:
Db        226 TLTQSVPMAGAVIVTTPQTVSLLDARRGLKMFQQMGVNVLGIVENMSYFIPPDLPDRQYD 285

Qy        243 ------GEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQA 296
                    ||        |:|          |||  :|  :  :         :|  |  .:|: |:
Db        286 LFGSGGGEKASKELNV------PLLGCVPLEIGLREGGDKGVPIVVSQPESASAKALTAI 339

Qy        297 AWRLVEQDSMKAVM 310
              | ::   : || |::
Db        340 AKQIAGKVSMAALV 353

RESULT 15
MRP_AQUAE
ID   MRP_AQUAE      STANDARD;      PRT;   364 AA.
AC   O66946;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mrp protein homolog.
GN   MRP OR AQ_737.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC   Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- SIMILARITY: BELONGS TO THE MRP/NBP35 FAMILY OF ATP BINDING
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE000705; AAC06915.1; -.
DR   InterPro; IPR002744; DUF59.
DR   InterPro; IPR000808; Mrp_family.
DR   Pfam; PF01883; DUF59; 1.
DR   PROSITE; PS01215; MRP; 1.
KW   ATP-binding; Complete proteome.
FT   NP_BIND     120    127       ATP (POTENTIAL).
SQ   SEQUENCE   364 AA;  39492 MW;  320AF95C191A272A CRC64;

  Query Match             12.8%;  Score 211;  DB 1;  Length 364;
  Best Local Similarity   27.1%;  Pred. No. 1.2e-09;
  Matches   75;  Conservative   50;  Mismatches  110;  Indels   42;  Gaps   10;

Qy         35 PSRRSPIRSVLQFNRKPELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAID 94
                |  . |.          | |  |  .: |     |:  .  |||||||:|  || :  ::|:: |: |  :|
Db         96 PQTQQPM-----FTRK-KVPG-VKHIIAVGSGKGGVGKSTVAANLAVALSQLGYKVGLLD 148

Qy         95 ADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMG 154
              ||:   ::  | ||:              |  :  |  :  .:|          .      |      |
Db        149 ADVYGPSVPTLFGLKG-------ERVTVD-QFQRIIPVEKYGLKILSIGFMLPSEDTPII 200

Qy        155 FGGKAL-EWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAI--TPANEAVLVTTPDITAL 211
               :  |   |  :  |  :   |:  ||:::|  |  |      ||          ||:||||    ||
Db        201 WRGPMLMKALTEFLFSTKWGNLDFLVMDLPPGTGDVQITLAQNVELTGAVVVTTPQDVAL 260

Qy        212 RDADRVTGLLECDGIRDIKMIVNRVRTDMI---------------KGEDMMSVLDVQEML 256
                |  :   :          |. : |   |  |   .|                      ||:      |  :
Db        261 ADVKKAVSM-----FREVNIPVLGVIENMAYFICPSDKQKYYIFGKGK----VAEFANAY 311

Qy        257 GLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAF 293
               || :|| || | ||    .::| |::|:: |  :    ||
Db        312 GLKILGSIPIDPEVAEKSDKGEPIVISHPDSEVAKAF 348

Search completed: April 16, 2003, 09:05:24
Job time : 28 secs
```

GenCore version 5 1 4_p5_4578

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:04:57 ; Search time 15 Seconds
(without alignments)
639.458 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLPLFSTNHQSLLLPSSL . . KAVMVEEEPKKPGPFSPPGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	144	8.7	359	4	US-09-134-001C-4842	Sequence 4842, Ap
2	88.5	5.4	264	3	US-08-969-644-16	Sequence 16, Appl
3	88.5	5.4	264	3	US-08-444-189-16	Sequence 16, Appl
4	88.5	5.4	264	4	US-08-468-544-16	Sequence 16, Appl
5	88.5	5.4	519	3	US-08-997-445D-2	Sequence 2, Appli
6	88	5.3	376	4	US-09-461-474-6	Sequence 6, Appli
7	88	5.3	732	5	PCT-US95-17026-2	Sequence 2, Appli
8	87	5.3	249	1	US-08-597-236-5	Sequence 5, Appli
9	87	5.3	249	1	US-08-746-682A-5	Sequence 5, Appli
10	86.5	5.2	416	3	US-09-320-878-18	Sequence 18, Appl
11	86.5	5.2	416	4	US-09-105-537-39	Sequence 39, Appl
12	86	5.2	410	4	US-09-134-001C-3927	Sequence 3927, Ap
13	85.5	5.2	1068	4	US-09-085-199B-11	Sequence 11, Appl
14	83.5	5.0	3077	6	5223423-2	Patent No. 5223423
15	83.5	5.0	3457	2	US-08-416-603-4	Sequence 4, Appli
16	82.5	5.0	367	2	US-08-896-320-1	Sequence 1, Appli
17	82.5	5.0	367	2	US-08-896-320-3	Sequence 3, Appli
18	82.5	5.0	396	4	US-09-461-474-2	Sequence 2, Appli
19	82.5	5.0	670	4	US-08-980-080-2	Sequence 2, Appli
20	82.5	5.0	756	1	US-08-434-730-16	Sequence 16, Appl
21	82	5.0	258	4	US-09-504-358-12	Sequence 12, Appl
22	82	5.0	258	4	US-09-954-314-12	Sequence 12, Appl
23	81.5	4.9	5087	4	US-09-144-085-1	Sequence 1, Appli
24	81	4.9	385	4	US-09-570-778A-12	Sequence 12, Appl
25	81	4.9	920	3	US-08-930-996A-8	Sequence 8, Appli
26	81	4.9	1705	4	US-08-669-785-4	Sequence 4, Appli
27	80.5	4.9	332	4	US-09-172-952-11	Sequence 11, Appl
28	80.5	4.9	812	1	US-08-446-794A-2	Sequence 2, Appli
29	80.5	4.9	812	1	US-08-750-007-3	Sequence 3, Appli
30	80.5	4.9	812	2	US-08-945-024-2	Sequence 2, Appli
31	79	4.8	1041	1	US-08-220-151-4	Sequence 4, Appli
32	79	4.8	1041	1	US-08-413-118-4	Sequence 4, Appli
33	79	4.8	1041	3	US-08-473-446-4	Sequence 4, Appli
34	78.5	4.7	371	2	US-08-837-593-8	Sequence 8, Appli
35	78.5	4.7	676	4	US-09-085-199B-7	Sequence 7, Appli
36	78.5	4.7	1186	4	US-09-134-001C-4522	Sequence 4522, Ap
37	78	4.7	332	4	US-09-461-474-15	Sequence 15, Appl
38	78	4.7	333	2	US-08-674-149A-2	Sequence 2, Appli
39	78	4.7	937	1	US-08-253-155A-31	Sequence 31, Appl
40	78	4.7	3567	2	US-07-642-734C-4	Sequence 4, Appli
41	78	4.7	3567	3	US-08-439-009A-4	Sequence 4, Appli
42	77.5	4.7	325	4	US-09-088-435-1	Sequence 1, Appli
43	77.5	4.7	338	4	US-09-134-001C-4795	Sequence 4795, Ap
44	77.5	4.7	2037	4	US-09-306-998-3	Sequence 3, Appli
45	77	4.7	420	4	US-09-066-047-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-134-001C-4842
; Sequence 4842, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4842
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4842

Query Match 8.7%; Score 144; DB 4; Length 359;
Best Local Similarity 24.3%; Pred. No. 9.3e-08;
Matches 55; Conservative 44; Mismatches 79; Indels 48; Gaps 8;

```
Qy      16 LPSSLSQKTLISSPRFVNNPSRR-SPIRSVLQFNRKPELAGETPRIVVITSGKGGVGKTT 74
           ||| . .      |.:  | :  |  :|  |   |       : | ||||||||:|
Db      85 LPSEVVE-------RYIGKGSEKPKTIEELLSQNNPVE-------FISIASGKGGVGKST 130

Qy      75 TTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI----NGDCRLDQAL 130
             |: ::|:| |  |  :|||:  ..  :.|  |      |:|    :|   :  |
Db     131 VAVNLAVALAREGKKVGLVDADIYGFSVPDMMGIDERPGIDGKEIIPVERHGVKVISMAF 190

Qy     131 VRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPE-GSPDFIIIDCPAGID-- 187
           ::                |. . |  |  ::    |  : |  |::::| | |
Db     191 FVEE----------------NAPVIWRGPMLGKMLTNFFTEVQWGELDYLLLDLPPGTGDV 235

Qy     188 AGFITAITPANEAVLVTTPDITALRDADR----------VTGLLE 222
           |  : :: |::: ::||||  ||   | |          : |::|
Db     236 ALDVHSMLPSSKEIIVTTPHPTAAFVAARAGAMAKHTEHTILGVIE 281
```

RESULT 2
US-08-969-644-16
; Sequence 16, Application US/08969644
; Patent No. 6096519
; GENERAL INFORMATION:
; APPLICANT: Ratti, Giulio
; APPLICANT: Comanducci, Maurizio

```
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/969,644
;      FILING DATE:  13-NOV-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/467,152
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/661,820
;      FILING DATE:
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-969-644-16

  Query Match             5.4%;  Score 88.5;  DB 3;  Length 264;
  Best Local Similarity   24.0%;  Pred. No. 0.11;
  Matches   44;  Conservative   26;  Mismatches   78;  Indels   35;  Gaps    7;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARY GFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
           :|  | ||| |||| · ||| ·||·· |  |:  | |    ||   ||
Db       4 LVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLD-PQSNLSSGLG------------ 50

Qy     120 INGDCRLDQALVRDKRWSNFEL---LCISKPRS--KLPMGFGGKALEWLVDALKTRPEGS 174
           | ||  : |  ... |   :|  |  |    |  |  ·    |   ·  |  ·
Db      51 --ASVRSDQKGLHDIVYTSNDLKSIICETKKDSVDLIPASFSSEQFREL--DIHRGPSNN 106

Qy     175 ------------PDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLE 222
                       |  ||| |  :      |    :. .  ||:  :.    ..   |
Db     107 LKLFLNEYCAPFYDICIIDTPPSLGGLTKEAFVAGDKLIACLTPEPFSILGLQKIREFLS 166

Qy     223 CDG 225
           |
Db     167 SVG 169


RESULT 3
US-08-444-189-16
; Sequence 16, Application US/08444189
; Patent No. 6110705
;  GENERAL INFORMATION:
;    APPLICANT:  Ratti, Giulio
```

```
;    APPLICANT:  Comanducci, Maurizio
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/444,189
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/180,528
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/991,512
;      FILING DATE:
;      APPLICATION NUMBER:  US/07/661,820
;      FILING DATE:
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-444-189-16

  Query Match             5.4%;  Score 88.5;  DB 3;  Length 264;
  Best Local Similarity   24.0%;  Pred. No. 0.11;
  Matches   44;  Conservative   26;  Mismatches   78;  Indels   35;  Gaps    7;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARY GFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
           :|  | ||| |||| : ||| ·||·· |  |·  | |    ||   ||
Db       4 LVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLD-PQSNLSSGLG------------ 50

Qy     120 INGDCRLDQALVRDKRWSNFEL---LCISKPRS--KLPMGFGGKALEWLVDALKTPPEGS 174
           | ||  : |  .: ·|   ·|  | |    |  |  ·    |   ·  |  .
Db      51 --ASVRSDQKGLHDIVYTSNDLKSIICETKKDSVDLIPASFSSEQFREL--DIHRGPSNN 106

Qy     175 ------------PDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLE 222
                       |  ||| |  .      |    .. .  ||.  ..    ..   |
Db     107 LKLFLNEYCAPFYDICIIDTPPSLGGLTKEAFVAGDKLIACLTPEPFSILGLQKIREFLS 166

Qy     223 CDG 225
           |
Db     167 SVG 169


RESULT 4
US-08-468-544-16
; Sequence 16, Application US/08468544
```

```
; Patent No. 6248563
;  GENERAL INFORMATION:
;    APPLICANT:  Ratti, Giulio
;    APPLICANT:  Comanducci, Maurizio
;    APPLICANT:  Tecce, Mario F.
;    APPLICANT:  Giuliani, Marzia M.
;    TITLE OF INVENTION:  pCTD PLASMID ISOLATED FROM CHLAMYDIA
;    TITLE OF INVENTION:  TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY
;    TITLE OF INVENTION:  THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID
;    TITLE OF INVENTION:  GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH & BIRCH
;      STREET:  301 N. Washington Street
;      CITY:  Falls Church
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22046-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,544
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/661,820
;      FILING DATE:  28-FEB-1991
;      APPLICATION NUMBER:  IT MI 91A000314
;      FILING DATE:  07-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Svensson, Leonard R.
;      REGISTRATION NUMBER:  30,330
;      REFERENCE/DOCKET NUMBER:  1267-202P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-241-1300
;      TELEFAX:  703-241-2848
;      TELEX:  248345
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  264 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-468-544-16

  Query Match             5.4%;  Score 88.5;  DB 4;  Length 264;
  Best Local Similarity   24.0%;  Pred. No. 0.11;
  Matches   44;  Conservative   26;  Mismatches   78;  Indels   35;  Gaps    7;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARY-GFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
           :|  | ||| |||| : ||| :||:: |  |:  | |    ||   ||
Db       4 LVFCSFKGGTGKTTLSLNVGCNLAQFLGKKVLLADLD-PQSNLSSGLG------------ 50

Qy     120 INGDCRLDQALVRDKRWSNFEL---LCISKPPS--KLPMGFGGKALEWLVDALKTPPEGS 174
                 | ||  : |  ::: :|   :| :|  |   :|  |  :    |   :   |  :
Db      51 --ASVRSDQKGLHDIVYTSNDLKSIICETKKDSVDLIPASFSSEQFREL--DIHRGPSNN 106

Qy     175 ------------PDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLE 222
                       |  ||| |  :      |    :: :   ||:  ::    ::   |
Db     107 LKLFLNEYCAPFYDICIIDTPPSLGGLTKEAFVAGDKLIACLTPEPFSILGLQKIREFLS 166

Qy     223 CDG 225
           |
Db     167 SVG 169


RESULT 5
US-08-997-445D-2
; Sequence 2, Application US/08997445D
; Patent No. 6043342
;  GENERAL INFORMATION:
;    APPLICANT:  Kocher, Olivier N.
;    TITLE OF INVENTION:  PDZK1 Protein Containing PDZ
;    TITLE OF INVENTION:  Interaction Domains
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  David Prashker, Esq.
;      STREET:  P.O. Box 67
;      CITY:  Brookline
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02146
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.40 Mb storage
;      COMPUTER:  IBM PS/1
;      OPERATING SYSTEM:  MS DOS
;      SOFTWARE:  WordPerfect version 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/997,445D
;      FILING DATE:  December 23, 1997
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  David Prashker
;      REGISTRATION NUMBER:  29,693
;      REFERENCE/DOCKET NUMBER:  BIS-037
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 232-7509
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  519 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-997-445D-2

  Query Match             5.4%;  Score 88.5;  DB 3;  Length 519;
  Best Local Similarity   19.3%;  Pred. No. 0.33;
  Matches   53;  Conservative   38;  Mismatches   76;  Indels  107;  Gaps   12;

Qy      32 VNNPSRRSPIRSVLQFNRKP---ELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGF 88
           |:  : :  :   :|| |:    :|    |||| :  |  |                |||
Db     212 VDKETDKRHVEQKIQFKRETASLKLLPHQPRIVEMKKGSNG----------------YGF 255

Qy      89 SVVA-----------IDA-----DLGLRNLDLLLGLENRVNYTCVEVINGDCRLD----- 127
             : |           ||:     : ||:| ||::     ||   || :: |  ::
Db     256 YLRAGSEQKGQIIKDIDSGSPAEEAGLKNNDLVVA----VNGESVETLDHDSVVEMIRKG 311

Qy     128 ----QALVRDKRWSNFELLCISKP-----PSKLPMGFGGKALEWLVDALKTPPEGSPDFI 178
                || ||   |  |      |       :|| |            ::|  |  :|  :
Db     312 GDQTSLLVVDKETDNMYLRAHFSPFLYYQSQELFNG----------SVKEAPAPTPTSL 360

Qy     179 IIDCP-----------------------------AGIDAGFITAITPANEAVLVTTPDI 208
           :  |                             |:   ||  :      |     |:
Db     361 EVSSPPDTTEEVDHKPKLCRLAKGENGYGFHLNAIRGLPGSFIKEVQKGGPA------DL 414

Qy     209 TALRDADRVTGLLECDGIRDI----KMIVNRVRT 238
           | | |  .   :| :|:  :    . .|.|..:
Db     415 AGLEDEDVI---IEVNGVNVLDEPYEKVVDRIQS 445


RESULT 6
US-09-461-474-6
; Sequence 6, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Plant Metal Transporters
; FILE REFERENCE: BB1303 US NA
;  CURRENT APPLICATION NUMBER: US/09/461,474
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: 60/112,562
```

```
;  EARLIER FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
;   LENGTH: 376
;   TYPE: PRT
;   ORGANISM: Triticum aestivum
US-09-461-474-6

  Query Match             5.3%;  Score 88;  DB 4;  Length 376;
  Best Local Similarity   34.0%;  Pred. No. 0.22;
  Matches   33;  Conservative   15;  Mismatches   35;  Indels   14;  Gaps    5;

Qy       3 SLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSPRSPIFS---VLQFNRKPELAGETPR 59
           || |||  ||       | |:||   |     |   : : |    |: : :  |   : |
Db      12 SLLLFS-RHQ-------LRQRTL--QPHRTFRPKMSTAVISAEDALEPSLQSLLDQRSLR 61

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDAD 96
             . :  ||||||||||. .: : :|:   ||. |  |
Db      62 WIFV-GGKGGVGKTTTSCSLAIQLAKVRRSVLLISTD 97


RESULT 7
PCT-US95-17026-2
; Sequence 2, Application PC/TUS9517026
;  GENERAL INFORMATION:
;    APPLICANT:  ZymoGenetics, Inc.
;    APPLICANT:  1201 Eastlake Avenue East
;    APPLICANT:  Seattle
;    APPLICANT:  WA
;    APPLICANT:  USA
;    APPLICANT:  98102
;    TITLE OF INVENTION:  Bovine Factor XIII
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/17026
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parker, Gary E
;      REGISTRATION NUMBER:  31-648
;      REFERENCE/DOCKET NUMBER:  94-18PC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6673
;      TELEFAX:  206-442-6678
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  732 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-17026-2

  Query Match             5.3%;  Score 88;  DB 5;  Length 732;
  Best Local Similarity   24.2%;  Pred. No. 0.65;
  Matches   72;  Conservative   38;  Mismatches   94;  Indels   94;  Gaps   20;

Qy      34 NPSRRSPIRSVLQFNRKPE---LAGETPRIVVITSGKGGVGKTTTTANVGLSL------- 83
           || . | . |. . | | .   :||     :         || . | :| . |
Db     255 NPIKVSRVGSAM-INAKDDEGVIAGSWDNVYAY-----GVPPSAWTGSVDILLEYKSSQK 308

Qy      84 -ARYGFSVVAIDADLGLRNLDL-LLGLENRV-NYTCVEVINGDCRLD---------QA 129
             |||  |      |: |  |  ||:  ||  ||       . : :||
Db     309 PVRYGQCWVF----AGVFNTFLRCLGIPARVVTNYFSAHDNDANLQLDIFLEEDGNVNSK 364

Qy     130 LVRDKRWSNFELLCISK---PRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDCPAGI 186
           | :|  | |:  | : :    |  ||:|||    | |    . |: : | :    || :
Db     365 LTKDSVW-NYH--CWNEAWMTRPDLPVGFGG----WQV--VDSTPQENSDGMYRCGPASV 415

Qy     187 -----------DAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC----------- 223
                      || |: |    |  ::     :|| :|   |   |:
Db     416 QAIKHGHVCFQFDAPFVFA--EVNSDLVY----VTAKKDGTHVVEALDTTHIGKLIVTKE 469

Qy     224 ---DGIRDIKMIVNRVRTDMIK----GEDMMSVLDVQEMLG----LSLLGVIPEDSEV 270
               ||.:||        ||  |     |:   |:   |:     |:  ||:   |:|
Db     470 IGGDGMKDI--------TDTYKFQEGQEEERLALETAMMYGAKKALNTEGVLKSKSDV 519


RESULT 8
US-08-597-236-5
; Sequence 5, Application US/08597236
; Patent No. 5733765
;  GENERAL INFORMATION:
;    APPLICANT:  STINGELE, Franscesca
;    APPLICANT:  MOLLET, Beat
;    TITLE OF INVENTION:  LACTIC BACTERIA PRODUCING
;    TITLE OF INVENTION:  EXOPOLYSACCHARIDES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americans
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/597,236
;      FILING DATE:
;      CLASSIFICATION:  426
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 95201669.9
;      FILING DATE:  20-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fanucci A., Allan
;      REGISTRATION NUMBER:  30256
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-597-236-5

  Query Match             5.3%;  Score 87;  DB 1;  Length 249;
  Best Local Similarity   20.8%;  Pred. No. 0.15;
  Matches   48;  Conservative   40;  Mismatches   79;  Indels   64;  Gaps   10;

Qy      41 IRSVLQFNRKPE----------LAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFS 89
           . | . | . | |           :|     . | | : | ||   . |: :| |  |
Db       7 VKSKVDFAKKTEEYYNAIRTNIQFSGAQMKVIAISSVEAGEGKSMISVNLAISFASVGLR 66

Qy      90 VVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNF--------EL 141
           : |||:   ||                |:.|  . ::     |  |||         |
Db      67 TLLIDAE-TRN --------------SVLSGTFKSNEFY---KGLSNFLSGNADLNET 105

Qy     142 LC---IS----------KPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDC-PAG 185
           .|   ||          | | |        |..  .:      |::|||  | |
Db     106 ICQTDISGLDVIASGPVPPNPTSLLQ----NDNFRHLMEVARS----CYDYVIIDTPPVG 157
```

```
Qy    186 IDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRV 236
Db    158 LVIDAVIIAHQADASLLVTEAGKIKRRFVTKAVEQLVESGSQFLGVVLNKV 208


RESULT 9
US-08-746-682A-5
; Sequence 5, Application US/08746682A
; Patent No. 5786184
;  GENERAL INFORMATION:
;    APPLICANT:  STINGELE, Franscesca
;    APPLICANT:  MOLLET, Beat
;    TITLE OF INVENTION:  LACTIC BACTERIA PRODUCING
;    TITLE OF INVENTION:  EXOPOLYSACCHARIDES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americans
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/746,682A
;      FILING DATE:  14-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/597,236
;      FILING DATE:  20-JUN-1995
;      APPLICATION NUMBER:  EP 95201669.9
;      FILING DATE:  20-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fanucci A., Allan
;      REGISTRATION NUMBER:  30256
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-746-682A-5

  Query Match             5.3%;  Score 87;  DB 1;  Length 249;
  Best Local Similarity   20.8%;  Pred. No. 0.15;
  Matches   48;  Conservative   40;  Mismatches   79;  Indels   64;  Gaps   10;

Qy      41 IRSVLQFNRKPE-----------LAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFS 89
Db       7 VKSKVDFAKKTEEYYNAIPTNIQFSGAQMKVIAISSVEAGEGKSMISVNLAISFASVGLR 66

Qy      90 VVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNF--------EL 141
Db      67 TLLIDAE--TRN---------------SVLSGTFKSNEPY---KGLSNFLSGNADLNET 105

Qy     142 LC---IS------------KPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDC-PAG 185
Db     106 ICQTDISGLDVIASGPVPPNPTSLLQ----NDNFRHLMEVARS----CYDYVIIDTPPVG 157

Qy     186 IDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRV 236
Db     158 LVIDAVIIAHQADASLLVTEAGKIKRRFVTKAVEQLVESGSQFLGVVLNKV 208


RESULT 10
US-09-320-878-18
; Sequence 18, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
;  APPLICANT: ASHLEY, Gary
;  APPLICANT: BETLACH, Melanie C.
;  APPLICANT: BETLACH, Mary C.
;  APPLICANT: McDANIEL, Robert
;  APPLICANT: TANG, Li
;  TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
;  FILE REFERENCE: 3006220021120
;  CURRENT APPLICATION NUMBER: US/09/320,878A
;  CURRENT FILING DATE: 1999-05-27
;  EARLIER APPLICATION NUMBER: CIP OF 09/141,908
;  EARLIER FILING DATE: 1998-08-28
;  EARLIER APPLICATION NUMBER: CIP OF 09/073,538
;  EARLIER FILING DATE: 1998-05-06
;  EARLIER APPLICATION NUMBER: CIP OF 08/846,247
;  EARLIER FILING DATE: 1997-04-30
;  EARLIER APPLICATION NUMBER: 60/119,139
;  EARLIER FILING DATE: 1999-02-08
;  EARLIER APPLICATION NUMBER: 60/100,880
;  EARLIER FILING DATE: 1998-09-22
;  EARLIER APPLICATION NUMBER: 60/087,080
;  EARLIER FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
;   LENGTH: 416
;   TYPE: PRT
;   ORGANISM: Streptomyces venezuelae
US-09-320-878-18

  Query Match             5.2%;  Score 86.5;  DB 3;  Length 416;
  Best Local Similarity   22.2%;  Pred. No. 0.39;
  Matches   42;  Conservative   31;  Mismatches   73;  Indels   43;  Gaps    7;

Qy     125 RLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDCPA 184
Db     102 RLRKLVAREFTMRRVELL---RPRVQ-------EIVDGLVDAMLAAPDGRADLM------ 145

Qy     185 GIDAGFITAITPANEAVLVTTPDITALR-----------DADRVTGLLECDGIRDIKMIV 233
Db     146 -ESLAWPLPITVISELLGVPEPDRAAFRVWTDAFVFPDDPAQAQTAMAEMSGY--LSRLI 202

Qy     234 NRVRTDMIKGEDMMSV-----------LDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVL 282
Db     203 DSKRGQ--DGEDLLSALVRTSDEDGSRLTSEELLGMAHILLVAGHETTVNLIANGMYALL 260

Qy     283 NKPPTLAGL 291
Db     261 SHPDQLAAL 269


RESULT 11
US-09-105-537-39
; Sequence 39, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
;  APPLICANT: Sherman, D.H.
;  APPLICANT: Liu, H.
;  APPLICANT: Xue, Y.
;  APPLICANT: Zhao, L.
;  TITLE OF INVENTION: DNA encoding methymycin and pikromycin
;  FILE REFERENCE: 600.438US1
;  CURRENT APPLICATION NUMBER: US/09/105,537A
;  CURRENT FILING DATE: 1998-06-26
;  NUMBER OF SEQ ID NOS: 43
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
;   LENGTH: 416
;   TYPE: PRT
;   ORGANISM: Streptomyces venezuelae
US-09-105-537-39

  Query Match             5.2%;  Score 86.5;  DB 4;  Length 416;
  Best Local Similarity   22.2%;  Pred. No. 0.39;
```

```
  Matches   42;  Conservative   31;  Mismatches   73;  Indels   43;  Gaps    7;

Qy     125 PLDQALVPDKPWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTPPEGSPDFIIIDCPA 184
           || : : |:          |||   ·||  ·      ·  · :: ||||·   |:|  |  ·
Db     102 RLPKLVAPEFTMPPVELL---PPPVQ-------EIVDGLVDAMLAAPDGRADLM------ 145

Qy     185 GIDAGFITAITPANEAVLVTTPDITALR-----------DADRVTGLLECDGIRDIKMIV 233
                :   ||  :! ·  |  ||  | |            |   |  : |   |    :  ::
Db     146 -ESLAWPLPITVISELLGVPEPDRAAFPVWTDAFVFPDDPAQAQTAMAEMSGY--LSPLI 202

Qy     234 NRVPTDMIKGEDMMSV--------  LDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVL 282
             : |      ||| :!                 !   | ||:  : ::       ·      |    :|
Db     203 DSKPGQ -DGEDLLSALVRTSDEDGSPLTSEELLGMAHILLVAGHETTVNLIANGMYALL 260

Qy     283 NKPPTLAGL 291
            : |  || |
Db     261 SHPDQLAAL 269
```

RESULT 12
US-09-134-001C-3927
; Sequence 3927, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3927
;   LENGTH: 410
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3927

```
  Query Match             5.2%;  Score 86;  DB 4;  Length 410;
  Best Local Similarity   24.6%;  Pred. No. 0.43;
  Matches   67;  Conservative   37;  Mismatches   96;  Indels   72;  Gaps   15;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLL----GLENRVNYTC 116
            | · ||  | ||||   ·· :  | |  ·  |  |:   |:|  |     |  ·| ·
Db      16 VTVLSGYLGSGKTTLLNHILKN--PEGPRIAVIVNDMSEVNIDKDLVAQGGGLSPTDEKL 73

Qy     117 VEVING--DCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEG- 173
           ||: ||   |  |    |:|:             : :         ||        :|  :    |
Db      74 VELSNGCICCTLRDDLLRE----------VERLVHK-----GG------IDQIVIESTGI 112

Qy     174 ------SPDFIIIDCPAGIDAGFI------TAITPANEAVL-VTTPDITALRD------- 213
                   :  |  ||   |||   |           ·  ||  |   : · |: · | ||
Db     113 SEPVPVAQTFSYIDEELGIDLTSICRLDTMVTVVDANRFVNDIRSEDLLADRDESVDDED 172

Qy     214 ----ADRVTGLLE-CDGIRDIKMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDS 268
                || ·   ·| |:         ||:·|·:  |:|  |        : ·:|           |   :  ::
Db     173 ERTIADLLIDQVEFCD-----VMIINKI--DLISDEALEKLENV--------LRALQPEA 217

Qy     269 EVIRSTNRGFPL--VLNKPPTLAGLAFEQAAW 298
           ·:|·· |   |  |||          | | |
Db     218 KIIKTVNAKVELSDVLNTQLFDFEKASESAGW 249
```

RESULT 13
US-09-085-199B-11
; Sequence 11, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
;   APPLICANT: Hayden, Michael R.
;   APPLICANT: Hackam, Abigail
;   APPLICANT: Huq, A.H.M. Mahbubul
;   APPLICANT: Chopra, Vikramjit Singh
;   APPLICANT: Kalchman, Michael
;   TITLE OF INVENTION: Apoptosis Modulators That Interact with the
;   TITLE OF INVENTION: Huntington's Disease Gene
;   NUMBER OF SEQUENCES: 44
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Oppedahl & Larson
;   STREET: PO Box 5270
;   CITY: Frisco
;   STATE: CO
;   COUNTRY: USA
;   ZIP: 80443-5270
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: MS DOS 5.0
;   SOFTWARE: WordPerfect
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/085,199B
;   FILING DATE:
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Larson, Marina T.
;   REGISTRATION NUMBER: 32038
;   REFERENCE/DOCKET NUMBER: UBC P-013US2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (970) 668-2050
;   TELEFAX: (970) 668-2052
; INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1068
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: no
;   ORIGINAL SOURCE:
;   ORGANISM: mouse
;   FEATURE:
;   OTHER INFORMATION: Huntington-interacting protein -mHIP1a
US-09-085-199B-11

```
  Query Match             5.2%;  Score 85.5;  DB 4;  Length 1068;
  Best Local Similarity   20.7%;  Pred. No. 2.3;
  Matches   69;  Conservative   52;  Mismatches  147;  Indels   65;  Gaps   14;

Qy      12 QSLLLPSSLSQKTLISSPPFVNNPS-PRSPIRSVLQFN----------PKPELAGETPPI 60
              ::·|  |    ·|:·      ||  |·|··  ·||              |:  ||         :
Db     731 RALELVGQLQDQTVLRRAQ----PSLMRAPLQGILQLGQDLKPKSLDVPQEELGAMVDKE 786

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRN--------LDLLLGLENRV 112
            :  ||    :         :| ||·  |  |  ·· ·  · :|               · ||·     ·
Db     787 MAATS--AAIEDAVPPIEDMMSQAPHESSGVKLEVNEPILNSCTDLMKAIPLLVMTSTSL 844

Qy     113 NYTCVEVINGDCRLDQALVPDKPWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTPPE 172
                  ||    |        ·    :. ||·     |·   ||           :|·|      |       |  ·
Db     845 QKEIVESGRGAATQQEFYAKNSRWTE-GLISASK-----AVGWGATQLVESADKVVLHMG 898

Qy     173 GSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALPDADPVTGLLECDGIPDIKMI 232
                |:  |                ·|:  |  ||          :  |   :::     ::  |   |     |  :
Db     899 KYEELIV--CSHEI---------AASTAQLVAASKVKANKNSPHLSRLQECS--RTVNER 945

Qy     233 VNPVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEV-IPSTNPGFPLVLNKPPTLAGL 291
                |       |·        |  ·  ||||·         :  |  ·|              ||          ||
Db     946 AANVVASTKSGQEQIEDPDTMDFSGLSLIKLKKQEMETQVP---------VLELEKTL--- 994

Qy     292 AFEQAAWRLVEQDSMKAVMV-------EEEPKK 317
              |        || |          |:            ||||  :
Db     995 --EAERVRLGELRKQHYVLAGGMGTPSEEEPSR 1025
```

RESULT 14
5223423 2

```
;Patent No. 5223423
;    APPLICANT: FRANCHINI, GENOVEFFA;WONG-STAAL, FLOSSIE;
;GALLO, ROBERT
;    TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT
;HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/07/331,212
;      FILING DATE: 03-31-1989
;SEQ ID NO:2:
;      LENGTH: 3077
5223423-2

  Query Match              5.0%;  Score 83.5;  DB 6;  Length 3077;
  Best Local Similarity   21.5%;  Pred. No. 21;
  Matches   61;  Conservative   48;  Mismatches  122;  Indels   53;  Gaps   15;

Qy       57 TPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTC 116
Db     2037 TPLCVTMNCNASTESAVATTSPSGPDMINDTDPCIQLNNCSGLREEDM---VECQFNMTG 2093

Qy      117 VEVINGDCRLDQALVPDKPWSNFELLC---ISKPRSKLPMGFGGKAL--------EWLVD 165
Db     2094 LE-------LDKKKQYSETWYSKDVVCESDNSTDRKRCYMNHCNTSVITESCDKHYW--D 2144

Qy      166 ALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDG 225
Db     2145 AMRFRYCAPPGFVLLRCNDTNYSGF----EPNCSKVVAST--CTRMMETQPSTWL----G 2194

Qy      226 IRDIKMIVNRVRTD-----MIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRST---NRG 277
Db     2195 F-------NGTPAENPTYIYWHGRDNPTIISLNKYYNLTILCRRPENKTVVPITLMSGRR 2247

Qy      278 F--PLVLNKPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRG 319
Db     2248 FHSQKIINKKPRQAWCRF-KGEWREAMQE VKQTLV-KHPRYKG 2288


RESULT 15
US-08-416-603-4
; Sequence 4, Application US/08416603
; Patent No. 5866780
;  GENERAL INFORMATION:
;    APPLICANT:  Law, Marcus
;    APPLICANT:  Hebara, Ledare
;    APPLICANT:  Reddick, Bradford B.
;    TITLE OF INVENTION:  Maize Chlorotic Dwarf Virus Genome and
;    TITLE OF INVENTION:  Uses Therefor
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/416,603
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lloyd, Jeffrey
;      REGISTRATION NUMBER:  35,589
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3457 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-416-603-4

  Query Match              5.0%;  Score 83.5;  DB 2;  Length 3457;
  Best Local Similarity   25.1%;  Pred. No. 26;
  Matches   53;  Conservative   30;  Mismatches   87;  Indels   41;  Gaps   14;

Qy       18 SSLSQKTLISSPRFVNNPSRRSPIRSVL-QFNRKPE-LAGETPRIVVITSGKGGVGKTTT 75
Db     2888 TSAATKTSI----------RKSPIHGLVGEIPTEPSILHAHDPRL-----PKDKIGKWDP 2932

Qy       76 TANVGLSLARYGFSVVAIDADLGLRNLD----LLLGLENRVNYTCV---EV-INGDCRLD 127
Db     2933 VIEASM---KYGSRITPFPVDQILEVEDHLSKMLANCENSKNKRQVNNLEIGING---ID 2986

Qy      128 QALVPDKPWSNFELLCISK-PRSK-LPMGFGGKALEWLVDALKTRPEGSPDFIIIDCPAG 185
Db     2987 QS----DYWQQIEMDTSSGWPYAKRKPVGAAGK--KWLFEQDGTYPSGKPRYVFGD--AG 3038

Qy      186 IDAGFITAITPANEAVLVTTPDITALRDADR 216
Db     3039 LIESYNSMLGEAKQGISPTVVTIECAKDERR 3069


Search completed: April 16, 2003, 09 09.02
Job time : 21 secs
```

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 16, 2003, 08:59:57 ; Search time 39 Seconds
(without alignments)
1113.839 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLRLFSTNHQSLLLPSSL ... FAVMVEEEPPKPGFFSFFGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1654	100.0	326	21	AAG46531	Arabidopsis thalia
2	1654	100.0	326	23	AAG77959	Arabidopsis thalia
3	1654	100.0	343	21	AAG46530	Arabidopsis thalia
4	1648	99.6	326	21	AAG12922	Arabidopsis thalia
5	1648	99.6	344	21	AAG12921	Arabidopsis thalia
6	1175.5	71.1	295	23	AAG77960	Tagetes erecta TeM
7	888	53.7	174	21	AAG12923	Arabidopsis thalia
8	888	53.7	174	21	AAG46532	Arabidopsis thalia
9	545.5	33.0	266	23	ABB49459	Listeria monocytog
10	366	22.1	260	22	AAB96685	Putative P. abyssi
11	327.5	19.8	245	22	AAB96374	Putative P. abyssi
12	252.5	15.3	258	22	AAB96719	Putative P. abyssi
13	224	13.5	294	19	AAY10390	H. pylori ORF 06ep
14	224	13.5	297	18	AAW20738	H. pylori cytoplas
15	209	12.6	425	18	AAW21013	H. pylori cytoplas
16	206	12.5	388	21	AAB11657	A. vitis hypersens
17	202	12.2	300	22	AAB96246	Putative P. abyssi
18	198	12.0	412	18	AAW20443	H. pylori cytoplas
19	197.5	11.9	279	22	AAG93171	C glutamicum prote
20	196.5	11.9	295	22	AAB96618	Putative P. abyssi
21	190	11.5	390	22	AAG81141	Mycobacterium tube
22	183	11.1	313	21	AAG11242	Arabidopsis thalia
23	183	11.1	313	21	AAG43414	Arabidopsis thalia
24	178	10.8	289	22	AAB94219	Human protein sequ
25	176	10.6	288	21	AAB59022	Breast and ovarian
26	173.5	10.5	375	22	AAG90990	C glutamicum prote
27	171.5	10.4	271	22	AAB92570	Human protein sequ
28	165	10.0	115	21	AAB40393	Human ORFX ORF157
29	165	10.0	115	23	ABP03821	Human ORFX protein
30	164.5	9.9	478	22	AAG90129	C glutamicum prote
31	163	9.9	293	22	ABB58650	Drosophila melanog
32	163	9.9	339	22	AAU39817	Propionibacterium
33	161	9.7	260	22	ABB71330	Drosophila melanog
34	161	9.7	400	21	AAG22284	Arabidopsis thalia
35	161	9.7	400	21	AAG39877	Arabidopsis thalia
36	161	9.7	532	21	AAG22283	Arabidopsis thalia
37	161	9.7	532	21	AAG39876	Arabidopsis thalia
38	161	9.7	547	21	AAG39875	Arabidopsis thalia
39	161	9.7	548	21	AAG22282	Arabidopsis thalia
40	159.5	9.6	257	21	AAY75106	Neisseria meningit
41	158.5	9.6	257	21	AAY75107	Neisseria meningit
42	157.5	9.5	286	22	AAB96245	Putative P. abyssi
43	154.5	9.3	350	21	AAG41434	Arabidopsis thalia
44	152.5	9.2	296	21	AAG41435	Arabidopsis thalia
45	152	9.2	342	23	ABB47911	Listeria monocytog

ALIGNMENTS

RESULT 1
AAG46531
ID AAG46531 standard; Protein; 326 AA.
XX
AC AAG46531;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58550.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US 0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

```
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PP   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             100.0%;  Score 1654;  DB 21;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-174;
  Matches  326;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Qy     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
           |||||||||||||||||||||||||||
Db     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326


RESULT 2
AAG77959
ID   AAG77959 standard; Protein; 326 AA.
XX
AC   AAG77959;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Arabidopsis thaliana AtMinD1 protein.
XX
KW   MinD; AtMinD1; chloroplast; plastid; photosynthesis.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200181601-A2.
XX
PD   01-NOV-2001.
XX
PF   18-APR-2001; 2001WO US12634.
XX
PR   19-APR-2000; 2000US-0553431.
XX
PA   (UYNE-) UNIV NEVADA & COMMUNITY COLLEGE SYSTEM.
XX
PI   Osteryoung KW;
XX
DR   WPI; 2002-041414/05.
DR   N-PSDB; AAH77248.
XX
PT   Novel DNA sequence involved in regulating plastid division in plants
PT   useful for obtaining transgenic plants with novel phenotypes,
PT   characterised by alterations in plastid shape, size and/or number
XX
PS   Example; Fig 1; 35pp; English.
XX
CC   The sequence represents A. thaliana AtMinD1. The invention relates
CC   to a DNA sequence isolated from its native genome comprising a plant
CC   MinD gene that functions in regulating plastid division. When ectopically
CC   expressed, the gene alters the number, shape and/or size of chloroplasts
CC   and other types of plastids present in plant cells. A genetic construct
CC   comprising the gene of the invention in sense or antisense orientation is
CC   useful for altering the size, shape and/or number of plastids, in
CC   particular chloroplasts in plant cells. Alteration in the plastid size,
CC   shape and number improves productivity or increases vigour due to
CC   enhanced photosynthetic capacity and allows enhanced production of
CC   commercially important compounds that accumulate naturally or as a result
CC   of genetic engineering.
XX
SQ   Sequence   326 AA;

  Query Match             100.0%;  Score 1654;  DB 23;  Length 326;
  Best Local Similarity   100.0%;  Pred. No. 7.3e-174;
  Matches  326;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Qy     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
           |||||||||||||||||||||||||||
Db     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326


RESULT 3
AAG46530
ID   AAG46530 standard; Protein; 343 AA.
XX
AC   AAG46530;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 58549.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
```

OS Arabidopsis thaliana.
XX
PN EP1033405 A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07 SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1654; DB 21; Length 343;
Best Local Similarity 100.0%; Pred. No. 7.9e-174;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRPSPIRSVLQFNRKPELAGETPRI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 77

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLPNLDLLLGLENRVNYTCVEVI 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLPNLDLLLGLENRVNYTCVEVI 137

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     138 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 197

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNPVRTDM 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 257

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 317

Qy     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
          |||||||||||||||||||||||||||
Db     318 VEQDSMKAVMVEEEPKKRGFFSFFGG 343
```

RESULT 4

AAG12922
ID AAG12922 standard; Protein; 326 AA.
XX
AC AAG12922;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12222.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.6%; Score 1648; DB 21; Length 326;
Best Local Similarity 99.7%; Pred. No. 3.3e-173;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60
           |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db       1 MASLRLFSTNHQSLLLPSSLSQKPLISSPRFVNNPSRPSPIRSVLQFNRKPELAGETPRI 60

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240

Qy    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Qy    301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
          |||||||||||||||||||||||||||
Db    301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
```

RESULT 5
AAG12921
ID AAG12921 standard, Protein; 344 AA.
XX
AC AAG12921;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12221.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.6%; Score 1648; DB 21; Length 344;
Best Local Similarity 99.7%; Pred. No. 3.6e-173;
Matches 325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSPRSPIRSVLQFNRKPELAGETPRI 60
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db      19 MASLRLFSTNHQSLLLPSSLSQKPLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 78

Qy      61 VVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLPNLDLLLGLENPVNYTCVEVI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      79 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 138

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     139 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 198

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     199 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 258

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     259 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIPSTNRGFPLVLNKPPTLAGLAFEQAAWRL 318

Qy     301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
           ||||||||||||||||||||||||||
Db     319 VEQDSMKAVMVEEEPKKRGFFSFFGG 344
```

RESULT 6
AAG77960
ID AAG77960 standard; Protein; 295 AA.
XX
AC AAG77960;
XX
DT 01-FEB-2002 (first entry)
XX
DE Tagetes erecta TeMinD protein.
XX
KW African marigold; MinD; TeMinD; chloroplast; plastid;
KW photosynthesis.
XX
OS Tagetes erecta.
XX
PN WO200181601-A2.
XX
PD 01-NOV-2001.
XX
PF 18-APR-2001; 2001WO-US12634.
XX
PR 19-APR-2000; 2000US-0553431.
XX
PR 19-APR-2000; 2000US-0553431.
XX
PA (UYNE-) UNIV NEVADA & COMMUNITY COLLEGE SYSTEM.
XX
PI Osteryoung KW;
XX
DR WPI; 2002-041414/05.
DR N-PSDB; AAH77249.
XX
PT Novel DNA sequence involved in regulating plastid division in plants
PT useful for obtaining transgenic plants with novel phenotypes,
PT characterised by alterations in plastid shape, size and/or number
XX
PS Example; Page 34-35; 35pp; English.
XX
CC The sequence represents T. erecta TeMinD protein. The invention relates
CC to a DNA sequence isolated from its native genome comprising a plant
CC MinD gene that functions in regulating plastid division. When ectopically
CC expressed, the gene alters the number, shape and/or size of chloroplasts
CC and other types of plastids present in plant cells. A genetic construct
CC comprising the gene of the invention in sense or antisense orientation is
CC useful for altering the size, shape and/or number of plastids, in
CC particular chloroplasts in plant cells. Alteration in the plastid size,
CC shape and number improves productivity or increases vigour due to
CC enhanced photosynthetic capacity and allows enhanced production of

```
CC   commercially important compounds that accumulate naturally or as a result
CC   of genetic engineering.
XX
SQ   Sequence   295 AA;

  Query Match             71.1%;  Score 1175.5,  DB 23,  Length 295;
  Best Local Similarity   72.0%;  Pred. No. 5e-121;
  Matches  247;  Conservative   15;  Mismatches   16;  Indels   65;  Gaps    7;

Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVN-NPSRRS----------PIRSVLQ 46
           | |||              |:: :|: |  |   ||  ::           ||||||
Db       1 MTSLRF----------LTEPSLVCSSTFPTFNPLHKTLTKPTPKPYPKPPPIRSVLQ 47

Qy      47 FNRKPELAGETPPIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLL 106
           :||||||||:|||                          |||||||:||||||||
Db      48 YNRKPELAGDTPP----------------------------VVAIDADVGLRNLDLLL 77

Qy     107 GLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDA 166
           |||||||| |||:|||||||||||||||||||||||||||||:|||||| || ||
Db      78 GLENRVNYTVVEVLNGDCRLDQALVRDKRWSNFELLCISKPPSKLPLGFGGKALVWL-DA 136

Qy     167 LKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGI 226
           || | || ||||:||||||||||||||||||||||||||||||||||||||||||||||
Db     137 LKDRQEGCPDFILIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGI 196

Qy     227 RDIKMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDS---EVIRSTNRGFPLVLN 283
           |||||||||||||:|:|||||||||||||||||     |:   ||||||||||||||
Db     197 RDIKMIVNRVRTDLIRGEDMMSVLDVQEMLGLSLL----SDTRGFEVIRSTNRGFPLVLN 252

Qy     284 KPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
           |||||||||||||||||||||||||||||||||||||||||||
Db     253 KPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 295


RESULT 7
AAG12923
ID   AAG12923 standard; Protein; 174 AA.
XX
AC   AAG12923;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 12223.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
```

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.7%; Score 888; DB 21, Length 174;
Best Local Similarity 100.0%; Pred. No. 1.3e-89;
Matches 174; Conservative 0; Mismatches 0, Indels 0, Gaps 0;

```
Qy    153 MGPGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALR 212
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGPGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALR 60

Qy    213 DADRVTGLLECDGIPDIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIR 272
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 DADRVTGLLECDGIPDIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIR 120

Qy    273 STNRGFPLVLNKPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKPGFFSFPGG 326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 STNRGFPLVLNKPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKPGFFSFPGG 174
```

RESULT 8
AAG46532
ID AAG46532 standard, Protein; 174 AA.
XX
AC AAG46532;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58551.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP 2000.
XX
PF 25-FEB 2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US 0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 53.7%; Score 888; DB 21; Length 174;
Best Local Similarity 100.0%; Pred. No. 1.3e-89;
Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 153 MGFGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALR 212
Db 1 MGFGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALR 60

Qy 213 DADRVTGLLECDGIRDIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIR 272
Db 61 DADRVTGLLECDGIRDIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIR 120

Qy 273 STNRGFPLVLNKPPTLAGLAFEQAAWPLVEQDSMKAVMVEEEPKKPGFFSFFGG 326
Db 121 STNRGFPLVLNKPPTLAGLAFEQAAWPLVEQDSMKAVMVEEEPKKPGFFSFFGG 174

RESULT 9
ABB49459
ID ABB49459 standard; Protein; 266 AA
XX
AC ABB49459;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #2163.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR01118.
XX
PR 11-APR-2000; 2000FR-0004629.
XX
PA (INSP) INST PASTEUR.
XX
PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
DR WPI; 2002-010914/01.
XX
PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
XX
PS Claim 6; SEQ ID No 2164; 192pp; French.
XX
CC The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present sequence is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 266 AA;

Query Match 33.0%; Score 545.5; DB 23; Length 266;
Best Local Similarity 43.9%; Pred. No. 1.9e-51;
Matches 115; Conservative 59; Mismatches 79; Indels 9; Gaps 5;

Qy 61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
Db 5 IVITSGKGGVGKTTSTANLGTALALQGKKVCLIDMDIGLRNLDVVLGLENPIIYDLVDVV 64

Qy 121 NGDCRLDQALVRDKPWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTPPEGSPDFIII 180
Db 65 EGRCKIHQAMIKDKPFDDL-LFLLPAAQTTDKNAVSGEQMVDLINQL--PPD--YDFILI 119

Qy 181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNPVPTDM 240
Db 120 DCPAGIETGYKNAVAGADKAIVVTTPEISAVRDADRIIGLLEKEDIEPPKLIINRIRTQM 179

Qy 241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNKGFPLVLNKPPTLAGLAFEQAAWKL 300
Db 180 MMNGDVMDIDEITTHLSIELLGIIIDDDEVIRSSNSGDPVAM-LPNNRASQGYRNIARRI 238

Qy 301 VEQDSMKAVMVEEEPKKRGFFS 322
Db 239 LGE---SIPLMSIETKKAGFFA 257

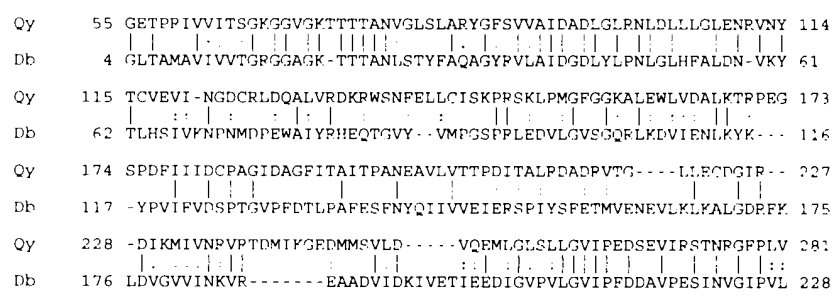
RESULT 10
AAB96685
ID AAB96685 standard; Protein; 260 AA.
XX
AC AAB96685;
XX
DT 29-OCT-2001 (first entry)
XX
DE Putative P. abyssi ATPase involved in chromosome partitioning #6.
XX
KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS Pyrococcus abyssi.
XX
PN FR2792651-A1.
XX
PD 27-OCT 2000.
XX
PF 21-APR-1999; 99FR-0005034.
XX
PR 21-APR-1999; 99FR-0005034.
XX

```
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 7; Pages 1443-1444; 1657pp; French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC   a hyperthermophilic archaeon, which is isolated from deep-sea
CC   hydrothermal vents. The present sequence is one such P. abyssi protein.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence   260 AA;

  Query Match             22.1%;  Score 366;  DB 22;  Length 260;
  Best Local Similarity   32.9%;  Pred. No  1.3e-31;
  Matches   94;  Conservative   54;  Mismatches   88;  Indels   50;  Gaps    7;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           |  :|  ||||| ||||| ||:|··||::|  |· |||:  : :| |:||:|:  :  |  :
Db       4 RSIVFASGKGGTGKTTTVANIGVALAQFGKDVILIDADITMANLSLILGMED-IPVTLHD 62

Qy     119 VINGDCRLDQALVPDKRWSNFELLCISKPRSKLPMGF----GGKALEWLVDALKTRPEG- 173
           |: |:  |  |:                  | |      || :|| :  |  |
Db      63 VLAGEADLKDAIYEG-----------------PAGVKVIPGGLSLEKIKKARAERLRDL 104

Qy     174 ------SPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIR 227
                 |||:|| |||::   |||:    | :·|| |·| |: |· ·   · |  |
Db     105 IREISQMGDFILIDAPAGLELTSITALLIGKELIIVTNPEIAAITDSLKTKLVAEKLGTL 164

Qy     228 DIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIPSTNPGFPLVLNKPPT 287
           ·  |·||| ··  ·    :|  ··· :| : :|| ·|| || |:·  | ||:  | :
Db     165 PLGAILNPVTSEKTE----LSPEEIEALLEVPVLGTVPEDPEVKRASAYGVPLVVKNPTS 220

Qy     288 LAGLAFEQAA-------WPLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
           | :|::: |       |:  | :|         | || | : |||
Db     221 PAAIAYKEIAAKLAGIKWKPPEPES-----  --PVYRIFPALFGG 257


RESULT 11
AAB96374
ID   AAB96374 standard; Protein; 245 AA.
XX
AC   AAB96374;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Putative P. abyssi ATPase involved in chromosome partitioning #3.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS   Pyrococcus abyssi.
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
```

```
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 7; Pages 1058-1059; 1657pp, French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC   a hyperthermophilic archaeon, which is isolated from deep-sea
CC   hydrothermal vents. The present sequence is one such P. abyssi protein.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence   245 AA;

  Query Match             19.8%;  Score 327.5;  DB 22;  Length 245;
  Best Local Similarity   34.6%;  Pred. No  2.1e-27;
  Matches   90;  Conservative   49;  Mismatches   90;  Indels   31;  Gaps    8;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           ||:  | |||||  |||| |||: ::|    |  |.|:|  ||  : || |.||::: |: |  :
Db       3 RIISIVSGKGGTGKTTVTANLSVALGEMGKKVLAVDGDLTMANLSLVLGVDD-VDVTLHD 61

Qy     119 VINGDCRLDQALVPDKRWSN----      FELLCISKPRSKLPMGFGGKALEWLVDALKT 169
           |: |: :|  |·     ·: |          ·| ·    || |||           ·· .||
Db      62 VLAGEAKLQDAIYM-TQFDNVYVLPGAVDWEHVMKADPR-KLPE---------VIKSLK- 109

Qy     170 RPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDI 229
             |  |||·|||||·    ::|·    ||·||! |·|: | |  ·|  ·|   |:  :
Db     110 ---GDYDFILIDCPAGLQLDAMSAMLSGEEALLVTNPEISCLTDTMKVGIVLRKAGLAIL 166

Qy     230 KMIVNRV-RTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTL 288
           |:||  ||:      |       |::: · || ||||  :   |  | | |  ||.
Db     167 GFILNRYGRTENDIPPDA-----AQDVMDVPLLAVIPEDPAIREGTLEGIPAVKYKPESE 221

Qy     289 AGLAFEQAAWRLVEQDSMKA 308
           || : |  : :  :||
Db     222 GAKAFVKLAQEVEKLAGIKA 241


RESULT 12
AAB96719
ID   AAB96719 standard; Protein; 258 AA.
XX
AC   AAB96719;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Putative P. abyssi ATPase involved in chromosome partitioning #7.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein.
XX
OS   Pyrococcus abyssi.
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
```

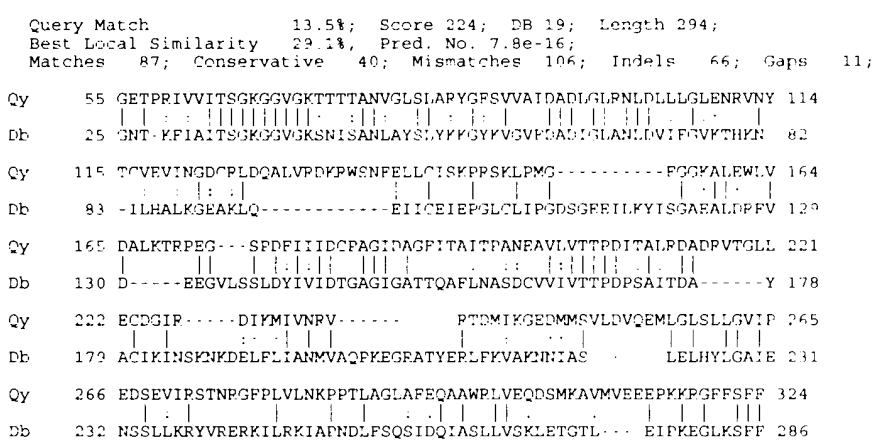
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
PI Querellou J, Weissenbach J, Saurin W, Heilig R;
XX
DR WPI; 2001-126236/14.
XX
PT New nucleotide sequences isolated from Pyrococcus abyssi encode
PT proteins useful in industry -
XX
PS Claim 7, Pages 1481 1482, 1657pp, French.
XX
CC The present invention relates to the genomic sequence of Pyrococcus
CC abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such P. abyssi protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade.
CC Note: This patent is in the same patent family as WO200065062, which
CC contains additional sequences as shown in AAB99132-AAB99143,
CC AAH75903-AAH75920 and AAG66436.
XX
SQ Sequence 258 AA;

Query Match 15.3%; Score 252.5; DB 22; Length 258;
Best Local Similarity 29.8%; Pred. No. 4.5e-19;
Matches 78; Conservative 57; Mismatches 98; Indels 29; Gaps 10;



RESULT 13
AAY10990
ID AAY10990 standard; Protein; 294 AA.
XX
AC AAY10990;
XX
DT 08-JUN-1999 (first entry)
XX
DE H. pylori ORF 06cp11202_4569693_c2_28 cellular protein.
XX
KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW secreted protein; cellular protein.
XX
OS Helicobacter pylori.
XX
PN WO9818323-A1.
XX
PD 07-MAY-1998.
XX
PF 28-OCT-1997; 97WO-US19575.
XX
PR 14-JUL-1997; 97US-0891928.
PR 28-OCT-1996; 96US-0739150.
PR 06-DEC-1996; 96US-0759739.
XX
PA (ASTR) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1998-271811/24.
DR N-PSDB; AAX30457.
XX
PT Helicobacter pylori nucleic acids and proteins - used to develop
PT products for the detection, prevention and treatment of H. pylori
PT infections
XX
PS Claims 27, 31; Page 195; 279pp; English.
XX
CC Recombinant or substantially pure preparations of H. pylori polypeptides
CC are disclosed, together with the nucleic acids encoding them. In all,
CC 73 ORFs are shown. The proteins are variously cell envelope proteins,
CC secreted proteins or other cellular proteins. Vaccines containing the
CC nucleic acids or proteins are claimed, as are probes containing at least
CC 8 nucleotides from the nucleic acid sequences. The vaccines are useful
CC for treating or reducing the risk of H. pylori infections, and the
CC probes can be used diagnostically for detecting the presence of
CC Helicobacter in a sample. The products are also of use in screening
CC for compounds having the ability to interfere with the H. pylori life
CC cycle or to inhibit H. pylori infection.
XX
SQ Sequence 294 AA;

Query Match 13.5%; Score 224; DB 19; Length 294;
Best Local Similarity 29.1%; Pred. No. 7.8e-16;
Matches 87; Conservative 40; Mismatches 106; Indels 66; Gaps 11;



RESULT 14
AAW20738
ID AAW20738 standard; protein; 297 AA.
XX
AC AAW20738;
XX
DT 16-JUL-1997 (first entry)
XX
DE H. pylori cytoplasmic protein, 06cp20302orf8.
XX
KW Cytoplasmic, vaccine, prevention, treatment, infection, identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
XX
OS Helicobacter pylori.
XX
PN WO9640893-A1.
XX

PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US09122.
XX
PR 01-APR-1996; 96US-0630405.
PR 07-JUN-1995; 95US-0487032.
XX
PA (ASTR) ASTRA AB.
XX
PI Berglindh OT, Smith D, Mellgaerd BL;
XX
DR WPI; 1997-052306/05.
DR N-PSDB; AAT67991.
XX
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX
PS Claim 61; Page 1154; 1481pp; English.
XX
CC The present sequence is a Helicobacter pylori cytoplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The genomic sequence of H. pylori (ATCC 55679) was determined from
CC overlapping contigs generated by mechanically shearing the bacterial
CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC and the predicted coding regions defined by computer evaluation. To
CC identify likely H. pylori antigens for vaccine development, the amino
CC acid sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
XX
SQ Sequence 297 AA;

Query Match 13.5%; Score 224; DB 18; Length 297;
Best Local Similarity 29.1%; Pred. No. 8e-16;
Matches 87; Conservative 40; Mismatches 106; Indels 66; Gaps 11;

```
Qy      55 GETPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNY 114
           | | : : |||||||||:  :||:  || : |: |   |||:|| |||:: |::   |
Db      28 GNT-KFIAITSGKGGVGKSNISANLAYSLYKKGYKVGVFDADIGLANLDVIFGVKTHKN- 85

Qy     115 TCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMG----------FGGKALEWLV 164
           .  . |: -|           |::|  :|   |  |              | :||.  |
Db      86 -ILHALKGEAKLQ-----------EIICEIEPGLCLIPGDSGEEILKYISGAEALDRFV 132

Qy     165 DALKTRPEG---SPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLL 221
           |      ||   | |:|:||  ||| |     :  ::  |:|||| :|: ||
Db     133 D-----EEGVLSSLDYIVIDTGAGIGATTQAFLNASDCVVIVTTPDPSAITDA------Y 181

Qy     222 ECDGIR--  --DIKMIVNRV  --------RTDMIKGEDMMSVLDVQEMLGLSLLGVIP 265
           |  |      :: :| | |           |   :   :: |       | |  || |
Db     182 ACIKINSKNKDELFLIANMVAQPKEGRATYERLFKVAKNNIAS-------LELHYLGAIE 234

Qy     266 EDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWPLVEQDSMKAVMVEEEPKKPGFFSFF 324
           | : |      |   | |  : :| | || :   :    | | | |||
Db     235 NSSLLKRYVRERKILRKIAPNDLFSQSIDQIASLLVSKLETGTL----EIPKEGLKSFF 289
```

RESULT 15
AAW21013
ID AAW21013 standard; protein; 425 AA.
XX
AC AAW21013;
XX
DT 22-JUL-1997 (first entry)
XX
DE H. pylori cytoplasmic protein, hp5e1521lorf13.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
XX
OS Helicobacter pylori.
XX
PN WO9640893-A1.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US09122.
XX
PR 01-APR-1996; 96US-0630405.
PR 07-JUN-1995; 95US-0487032.
XX
PA (ASTR) ASTRA AB.
XX
PI Berglindh OT, Smith D, Mellgaerd BL;
XX
DR WPI; 1997-052306/05.
DR N-PSDB; AAT68266.
XX
PT Helicobacter pylori nucleic acid sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX
PS Claim 61; Page 1402-1403; 1481pp; English.
XX
CC The present sequence is a Helicobacter pylori cytoplasmic protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds, useful
CC as potential H. pylori life cycle activators or inhibitors. The genomic
CC sequence of H. pylori (ATCC 55679) was determined from overlapping
CC contigs generated by mechanically shearing the bacterial DNA. The
CC sequences were analysed for ORF of at least 180 nucleotides, and the
CC predicted coding regions defined by computer evaluation. To identify
CC likely H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from H. pylori by PCR amplification for recombinant polypeptide
CC production, e.g. in E. coli hosts.
XX
SQ Sequence 425 AA;

Query Match 12.6%; Score 209; DB 18; Length 425;
Best Local Similarity 24.2%; Pred No. 6.4e-14;
Matches 80; Conservative 57; Mismatches 112; Indels 82; Gaps 12;

```
Qy       8 STNHQSLLLPSSLSQKTLISSPRFVNNPSPPSPIPSVLQFNPKPE--------LAGETPR 59
           |.   | :|  ..|:       : :|   :  |          ||:        ||
Db     105 SSEETSAILPENISKAMQEKGVKALNLDIKTPP---------KPQAPKPTTYNLAKNIKH 155

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
           :|.|.|||||||||:||: |: :.||     |  .|||:   |:  ::||:
Db     156 VVMISSGKGGVGKSTTSVNLSIALANLNQKVGLLDADVYGPNIPRMMGLQ---------- 205

Qy     120 INGDCRLD---QALVRDKRWSNFELLCISKPPSKLPMGF---GGKALEWLVDALKTRPE- 172
           | |  :|   . |: | :           | : ||     |::| |    |    |
Db     206 -NADVIMDPSGKKLIPLKAFG----------VSVMSMGLLYDEGQSLIWRGPMLMRAIEQ 254

Qy     173 -------GSPDFIIIDCPAGIDAGFIT--AITPANEAVLVTTPDITALRDADR------- 216
                 |  | :::| | |     :|     | :  : |||| | :| || |
Db     255 MLSDIIWGDLDVLVVDMPPGTGDAQLTLAQAVPLSAGITVTTPQIVSLDDAKRSLDMFKK 314

Qy     217 --- VTGLLE---  --CDGIRDIKMIVNFVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPE 266
               : |::|      |:  :         :   | | : || |   |     :|  :|
Db     315 LHIPIAGIVENMGSFVCEHCK---------KESEIFGSNSMSGL--LEAYNTQILAKLPL 363

Qy     267 DSEVIRSTNRGFPLVLNKPPTLAGLAFEQAA 297
           : :|    ..| |:|:: | :::   ||: |
Db     364 EPKVPLGGDKGEPIVISHPTSVSAKIFEKMA 394
```

Search completed: April 16, 2003, 09:06:09
Job time : 41 secs

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:04:32 ; Search Time 45 Seconds
(without alignments)
696.440 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLRLFSTNHQSLLLPSSL ..... KAVMVEEEPKKPGFFSFFGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	792	47.9	282	2	T07303	cell division inhi
2	657.5	39.8	268	2	AI2237	septum site-determ
3	643	38.9	266	2	S76970	hypothetical prote
4	573.5	34.7	276	2	D75478	septum site-determ
5	559.5	33.8	263	2	B97054	septum site-determ
6	546.5	33.0	266	2	AB1630	cell division inhi
7	545.5	33.0	266	2	AH1267	cell division inhi
8	536.5	32.4	264	2	C84028	septum site-determ
9	523	31.6	268	1	G45239	cell division inhi
10	516	31.2	262	2	F70375	cell division inhi
11	514.5	31.1	269	2	C82695	septum site-determ
12	506.5	30.6	268	2	C64561	cell division inhi
13	500.5	30.3	268	2	B71945	cell division inhi
14	487	29.4	270	2	AE0253	septum site-determ
15	483	29.2	270	2	AH0724	septum site determ
16	481.5	29.1	271	2	C98327	cell division inhi
17	481.5	29.1	271	2	AB2956	cell division inhi
18	480.5	29.1	271	2	F83239	cell division inhi
19	479	29.0	270	1	CCECID	cell division inhi
20	479	29.0	270	2	E90837	cell division inhi
21	479	29.0	270	2	F85695	cell division inhi
22	475.5	28.7	271	2	A72200	septum site-determ
23	468	28.3	276	2	C82136	septum site-determ
24	460	27.8	270	2	B84968	septum site-determ
25	452	27.3	271	2	A96011	probable cell divi
26	449.5	27.2	271	2	C81230	septum site-determ
27	388	23.5	245	2	AE3625	cell division inhi
28	366	22.1	260	2	F75175	cell division inhi
29	344.5	20.8	259	2	B69113	cell division inhi
30	341.5	20.6	245	2	C71105	probable cell divi
31	335	20.3	263	2	H69336	cell division inhi
32	327.5	19.8	245	2	A75056	cell division inhi
33	300	18.1	264	1	C64368	cell division inhi
34	266	16.1	252	2	F71191	hypothetical prote
35	262.5	15.9	305	2	A84363	cell division inhi
36	254	15.4	217	2	H71038	probable cell divi
37	249	15.1	280	2	B83463	flagellar synthesi
38	247.5	15.0	251	2	H75100	cell division inhi
39	246.5	14.9	288	2	D83954	hypothetical prote
40	243.5	14.7	323	2	B84251	cell division inhi
41	242	14.6	263	1	B64321	cell division inhi
42	241	14.6	288	2	C81422	probable ATP-bindi
43	239.5	14.5	304	2	D71290	probable ATP-bindi
44	238.5	14.4	295	2	E70133	ATP-binding protei
45	236	14.3	287	2	D97164	ATPases involved i

ALIGNMENTS

RESULT 1
T07303
cell division inhibitor - Chlorella vulgaris chloroplast
C;Species: chloroplast Chlorella vulgaris
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C;Accession: T07303
R;Wakasugi, T.; Nagai, T ; Kapoor, M ; Sugita, M.; Ito, M.; Ito, S., Tsudzuki, J., Nakash
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A;Reference number: Z15985, MUID:97303241; PMID:9159184
A;Accession: T07303
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-282 <WAK>
A;Cross-references: EMBL:AB001684; NID.g2224352, PIDN:BAA57951.1, PID.g2224467
C;Genetics:
A;Gene: minD
A;Genome: chloroplast
C;Superfamily: cell division inhibitor minD
C;Keywords: chloroplast

Query Match 47.9%; Score 792; DB 2; Length 282;
Best Local Similarity 60.4%; Pred. No. 5e-56;
Matches 165; Conservative 39; Mismatches 55; Indels 14, Gaps 5;

```
Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           |::|||||||||||||||:|:|:|| |: | ||||:|||||||||||||| || ::
Db      17 RVIVITSGKGGVGKTTTTANLGMSIARLGYRVALIDADIGLRNLDLLLGLENRVLYTAMD 76

Qy     119 VINGDCPLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
           :: | |||||||:|||| |  || ||| | |       | :: |:|::|   |    |:
Db      77 IVEGQCRLDQALIRDKRWKNLALLAISKNRQK--YNVTRKNMQNLIDSVK---ELGFQFV 131

Qy     179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEQDGIPDIKMIVNRVRT 238
           :|||||||| ||| ||  | |||:||||:|||:|||||| |||| :|| ::|::|||||
Db     132 LIDCPAGIDVGFINAIASAQEAVIVTTPEITAIRDADRVAGLLEANGIYNVKLLVNRVRP 191

Qy     239 DMIKGRDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
           |||:  ||||| |||||||: ||| |||||: || |||:| ||||||  ||:|:||| ||
Db     192 DMIQKNDMMSVPDVQEMLGIPLLGAIPEDTSVIISTNKGEPLVLNKKLTLSGIAFENAAR 251

Qy     299 RLV-EQDSMKAVMVEEEPKKRGFF----SFFGG 326
           ||: :||      ::    ::| |     || |
Db     252 PLIGKQD----YPIDLTSPQKGMPQKLQEFFLG 280
```

RESULT 2
AI2237
septum site-determining protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C;Accession: AI2237
R;Kaneko, T.; Nakamura, Y.; Wolk, C P ; Kuritz, T.; Sasamoto, S.; Watanabe, A ; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M ; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2237
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <KUP>
A;Cross-references: GB:BA000019; PIDN:BAB75155.1; PID:g17132589; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

Query Match 39.8%; Score 657.5; DB 2; Length 268;
Best Local Similarity 54.4%; Pred. No. 2.8e-45;
Matches 137; Conservative 48; Mismatches 54; Indels 13; Gaps 4;

Qy 59 RIVVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLRNLDLLLGLENPVNYTCVE 118
Db 3 RIIVITSGKGGVGKTTVSANLGMALAKMGRQVALVDADFGLRNLDLLLGLENRIVYTAVE 62

Qy 119 VINGDCRLDQALVPDKPWSNFELLCISKPPSK---LPMGFGGKALEWLVDALKTPPEGSP 175
Db 63 VLSRECRLEQALVKDKRQPNLVLLPAAQNRTKDAVTP-----EQMKLLVNALAQKYQ--- 114

Qy 176 DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNR 235
Db 115 -YVIIDSPAGIENGFKNAIAPAKEALIVTTPEISAVRDADRVVGLLEAQGVKPVHLIINR 173

Qy 236 VRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQ 295
Db 174 IRPAMVRANDMMSVQDVQELLAIPLIGVLPDDERVIVSTNRGEPLVLGDTPSLAAVAVEN 233

Qy 296 AAWRLVEQDSMK 307
Db 234 IAPPL-EGDTVE 244

RESULT 3
S76970
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S76970
R;Kaneko, T ; Sato, S ; Kotani, H.; Tanaka, A.; Asamizu, E., Nakamura, Y., Miyajima, N.;
O, K ; Okumura, S ; Shimpo, S ; Takeuchi, C ; Wada, T.; Watanabe, A ; Yamada, M., Yasuda
DNA Res 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322, MUID 97061201, PMID:8905231
A;Accession: S76970
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <KAN>
A;Cross references: EMBL:D64005; GR:AB001339; NID:g1001779; PIDN:BAA10662.1; PID:g100178
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: cell division inhibitor minD

Query Match 38.9%; Score 643; DB 2; Length 266;
Best Local Similarity 52.8%; Pred. No. 4.1e-44;
Matches 130; Conservative 51; Mismatches 59; Indels 6; Gaps 2;

Qy 59 RIVVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLRNLDLLLGLENPVNYTCVE 118
Db 3 RIIVVTSGKGGVGKTTTTANLGAALAPLGKKVVLIDADFGLRNLDLLLGLEQPIVYTAID 62

Qy 119 VINGDCRLDQALVPDKPWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTPPEGSPDFI 178
Db 63 VLADECTIDKALVKDKPLPNLVLLPAAQNRSK--DAINAEQMQSLVEQLKDK----FDYI 116

Qy 179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALPDADRVTGLLECDGIRDIKMIVNPVRT 238
Db 117 IIDCPAGIEAGFRNAVAPAQEAIIVTTPEMSAVRDADRVIGLLEAEDIGKISLIVNRLRP 176

Qy 239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
Db 177 EMVQLNQMISVEDILDLLAVPLIGILPDDQKIIISTNKGEPLVMEEPLSVPGLAFQNIAR 236

Qy 299 RLVEQD 304
Db 237 RLEGQD 242

RESULT 4
D75478
septum site-determining protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: D75478
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; D
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mal
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75478
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-276 <WHI>
A;Cross-references: GR:AE001931; GR:AE000513; NID:g6458461; PIDN:AAF10331.1; PID:g645846
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR0752
A;Map position: 1
C;Superfamily: cell division inhibitor minD

Query Match 34.7%; Score 573.5; DB 2; Length 276;
Best Local Similarity 44.3%; Pred. No. 1.6e-38;
Matches 125; Conservative 52; Mismatches 72; Indels 33; Gaps 5;

Qy 59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
Db 12 KVIVVTSGKGGVGKTTTTANIGAALAPLGEKVVVIDVDVGLPNLDVVMGLESPVVFDLVD 71

Qy 119 VINGDCPLDQALVPDKPWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTP-------- 170
Db 72 VLEGKCRMNQALIPDKRVENLHLLPASQTRDK-------------DALDPEVPKEVVKG 117

Qy 171 PEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRD 228
Db 118 LLEEEGFDRVLIDSPAGIESGFRTAAAPAEGALVVVNPEVSSVRDADRIIGLLEAQQITE 177

Qy 229 IKMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTL 288
Db 178 IPLVVNPLRPKMVASGNMLSIDDMVDILGVKPIGIVPEDEGIIVSTNVGEPAVLGK-TK 235

Qy 289 AGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFS----FFGG 326
Db 236 AGDAFMATAQRIQGQDVPFPKLTEEE---KGIWAAIRRLFGG 274

RESULT 5
B97054
septum site-determining protein MinD, ATPase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30 Sep 2001
C;Accession: B97054
R;Nelling, J ; Breton, G., Omelchenko, M.V ; Markarova, K.S ; Zeng, Q., Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clos
A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: B97054
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79221.1; PID:g15024176; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1249
C;Superfamily: cell division inhibitor minD

Query Match 33.8%; Score 559.5; DB 2; Length 263;
Best Local Similarity 44.7%; Pred. No. 2e-37;
Matches 117; Conservative 57; Mismatches 77; Indels 11; Gaps 4;

Qy 61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
Db 5 IVITSGKGGVGKTTTTANIGTALAAMKKKVVLVDGDTGLRNLDVLMGLENRIVFTLLDVV 64

Qy 121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
Db 65 EGNCRLKQALIKDKHYENLALLPTAQTRDKNDV--KPEQMLKLVNELKE----EFDYVII 118

Qy 181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
Db 119 DCPAGIEQGFENAVIGADRAIVVVNPEVTSVRDADRVIGKIDAKGIEDHQVIVNRIDYEM 178

Qy 241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
Db 179 VKRGDMLGIEDVIDNLAIKLIGVVPNDKQITVSTNKGEPIVLNQNAN-AGKAFRDIARRV 237

Qy 301 VEQDSMKAVMVEEEPKKRGFFS 322
Db 238 LGEE----VPFEKYETQTGFIA 255

RESULT 6
AB1630
cell division inhibitor (septum placement) protein MinD homolog minD [imported] - Lister
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AB1630
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1630
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96810.1; PID:g16414066; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

Query Match 33.0%; Score 546.5; DB 2; Length 266;
Best Local Similarity 43.3%; Pred. No. 2.2e-36;
Matches 117; Conservative 60; Mismatches 80; Indels 13; Gaps 6;

Qy 61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
Db 5 IVITSGKGGVGKTTSTANLGTALALQGKKVCLIDMDIGLRNLDVVLGLENRIIYDLVDVV 64

Qy 121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
Db 65 EGRCKIHQAMIKDKRFDDL-LFLLPAAQTTDKNAVSGEQMIELINQL--RPD--YDYILI 119

Qy 181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
Db 120 DCPAGIETGYKNAVAGADKAIVVTTPEISAVRDADRIIGLLEKEDIEPPKLIINRIPTQM 179

Qy 241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
Db 180 MVNGDVMDIDEITTHLSIELLGIIIDDDEVIFSSNSGDPVAM LPNNRASQGYRNIARRI 238

Qy 301 VEQDSMKAVMVEEEPKKRGFFS----FFGG 326
Db 239 LGE---SIPLMSIEAKKPGFFTRLKQLFGG 265

RESULT 7
AH1267
cell division inhibitor (septum placement) protein MinD homolog minD [imported] - Listeri
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AH1267
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1267
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99622.1; PID:g16410973; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

Query Match 33.0%; Score 545.5; DB 2; Length 266;
Best Local Similarity 43.9%; Pred. No. 2.6e-36;
Matches 115; Conservative 59; Mismatches 79; Indels 9; Gaps 5;

Qy 61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
Db 5 IVITSGKGGVGKTTSTANLGTALALQGKKVCLIDMDIGLRNLDVVLGLENRIIYDLVDVV 64

Qy 121 NGDCRLDQALVRDKRWSNFELLCISKPPSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
Db 65 EGRCKIHQAMIKDKRFDDL-LFLLPAAQTTDKNAVSGEQMVDLINQL--RPD--YDFILI 119

Qy 181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
Db 120 DCPAGIETGYKNAVAGADKAIVVTTPEISAVRDADRIIGLLEKRDIEPPKLIINRIPTQM 179

Qy 241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
Db 180 MVNGDVMDIDEITTHLSIELLGIIIDDDEVIFSSNSGDPVAM-LPNNRASQGYRNIARRI 238

Qy 301 VEQDSMKAVMVEEEPKKRGFFS 322
Db 239 LGE---SIPLMSIETKKAGFFA 257

RESULT 8
C84028
septum site-determining protein minD [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: C84028
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C84028
A;Status: preliminary
A;Molecule type: DNA

```
A;Residues: 1-264 <STO>
A;Cross-references: GB:AP001517; GB:BA000004; NID:g10175500; PIDN:BAB06746.1; GSPDB:GN0
A;Experimental source: strain C 125
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

  Query Match             32.4%;  Score 536.5;  DB 2;  Length 264;
  Best Local Similarity   41.1%;  Pred. No. 1.4e-35;
  Matches  109;  Conservative   66;  Mismatches   83;  Indels    7;  Gaps    3;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
              :||||||||||||||:||:| :||  |  |  :| |:||||||:::|||||: |  |:|:
Db       5 IVITSGKGGVGKTTTSANIGTALALSGKKVCLVDTDIGLRNLDVVMGLENRIIYDLVDVV 64

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
              | ||| |||::|||:     ||   :: : |       : :: :|: ||      |:::|
Db      65 EGRCRLKQALIKDKRFECLNLLPAAQTKDK--SAVTPEQMKEIVEELKQ----EYDYVLI 118

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
              ||||||: ||  |.  |:.|..||||.|:::|||||: ||||  :    ::.|||:|  |
Db     119 DCPAGIEQGFKNAVAGADKAIVVTTPEISSVRDADRIIGLLEKEEVEAPRLVVNRIRGHM 178

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
              :|  .|. |  ..   .| : |||:: :|  ||. .|:| |: |: | : | |:   | |:
Db     179 MKNGEMLDVDEIVSILAIELLGIVVDDENVIKFSNKGEPIALH-FDSKASVAYRNIARRI 237

Qy     301 VEQDSMKAVMVEEEPKKRGFFSFFG 325
              : :          :|:        ||||
Db     238 LGETVPLMSFEQEKGVLAKIKSFFG 262


RESULT 9
G45239
cell division inhibitor minD [validated] - Bacillus subtilis
N;Alternate names: septum placement determinant minD
C;Species: Bacillus subtilis
C;Date: 10-Jun-1993 #sequence_revision 04-Oct-1996 #text_change 19-Jan-2001
C;Accession: S31205; G45239; F45240; A69658; S27521; S29866
R;Lee, S.; Price, C.W.
Mol. Microbiol. 7, 601-610, 1993
A;Title: The minCD locus of Bacillus subtilis lacks the minE determinant that provides t
A;Reference number: S31204; MUID:93211302; PMID:8459776
A;Accession: S31205
A;Molecule type: DNA
A;Residues: 1-268 <LEE>
A;Cross-references: EMBL:Z15113; NID:g49307; PIDN:CAA78818.1; PID:g580893
R;Levin, P.A.; Margolis, P.S.; Setlow, P.; Losick, R.; Sun, D.
J. Bacteriol. 174, 6717-6728, 1992
A;Title: Identification of Bacillus subtilis genes for septum placement and shape determ
A;Reference number: A45239; MUID:93015731; PMID:1400224
A;Accession: G45239
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-268 <LEV>
A;Cross-references: GB:M96343; NID:g142852; PIDN:AAA22401.1; PID:g142859
A;Note: sequence extracted from NCBI backbone (NCBIP:116563)
R;Varley, A.W.; Stewart, G.C.
J. Bacteriol. 174, 6729-6742, 1992
A;Title: The divIVB region of the Bacillus subtilis chromosome encodes homologs of Esche
A;Reference number: A45240; MUID:93015732; PMID:1400225
A;Accession: F45240
A;Molecule type: DNA
A;Residues: 1-268 <VA2>
A;Cross-references: EMBL:M95582; NID:g143211; PIDN:AAA22609.1; PID:g143216
A;Note: sequence extracted from NCBI backbone (NCBIP:116574)
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
```

```
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69658
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-268 <KUN>
A;Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14759.1; PID:g2635264
A;Experimental source: strain 168
C;Genetics:
A;Gene: minD
A;Start codon: TTG
C;Function:
A;Description: mediates, together with minC, cell division site selection by specifically
C;Superfamily: cell division inhibitor minD
C;Keywords: cell division control; nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;16/Binding site: ATP (Lys) #status predicted

  Query Match             31.6%;  Score 523;  DB 1;  Length 268;
  Best Local Similarity   41.1%;  Pred. No. 1.7e-34;
  Matches  111;  Conservative   68;  Mismatches   77;  Indels   14;  Gaps    7;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
              :||||||||||||||:||:| .|| |  |  .| |.||||||:::|||||: |  |.|:
Db       5 IVITSGKGGVGKTTTSANLGTALAILGKRVCLVDTDIGLRNLDVVMGLENRIIYDLVDVV 64

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
              | |:: ||||:|||: .  |  :           : : :|  ||        |:.||
Db      65 EGRCKMHQALVKDKRFDDL-LYLMPAAQTSDKTAVAPEQIKNMVQELKQ----EFDYVII 119

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC-DGIRDIKMIVNPVPTD 239
              ||||||. |.  |..  |:.|::||||:|:|:||||: ||||  :    ::.||| |
Db     120 DCPAGIEQGYKNAVSGADKAIVVTTPEISAVRDADRIIGLLEQEENIEPPRLVVNRIRNH 179

Qy     240 MIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWR 299
              ::|  | | : :: : | : |||:: :| ||:...| | |: :: |   | .|. | |
Db     180 LMKNGDTMDIDEIVQHLSIDLLGIVADDDEVIKASNHGEPIAMD-PKNPASIAYRNIAPP 238

Qy     300 LVEQDSMKAVMVEEEPKKRGFF----SFFG 325
              :: :|:   ::||: |  |        ||||
Db     239 IL-GESVPLQVLEEQNK--GMMAKIKSFFG 265


RESULT 10
F70375
cell division inhibitor (septum placement) minD2 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 26-Aug-1999
C;Accession: F70375
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: F70375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <AQF>
A;Cross-references: GB:AE000712; NID:g2983411; PIDN:AAC06996.1; PID:g2983413; GP:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: minD2
C;Superfamily: cell division inhibitor minD

  Query Match             31.2%;  Score 516;  DB 2;  Length 262;
  Best Local Similarity   41.4%;  Pred. No. 6e-34;
```

```
  Matches 108;  Conservative  65;  Mismatches  70;  Indels  18;  Gaps   3;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
Db       4 VIVITSGKGGVGKTTLTANIGTALAKLGKKVLLIDADIGLRNLDMILGLENRIVYDILDV 63

Qy     120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
Db      64 LEGRVPYEKALVKDKRGLSLWLLPANQRANKDVIDIE----KWNKTVEEIKNSGNYDYIL 119

Qy     180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTD 239
Db     120 VDSPAGIEKGFQIAVSPADKALIVVNPEVSSIRDADRVIGLLESMDKRNYKVIVNRIKWE 179

Qy     240 MIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNK--PPTLAGLAFEQAA 297
Db     180 MVKRGAMLSVEDIVDILKAEIIGIIPEEPKLVDFTNPGEPIVLDEKFPASQAII----  233

Qy     298 WRLVEQDSMKAVMVEEEPKKR 318
Db     234 ------DTARRLMGESIPLKR 248
```

RESULT 11
C82695
septum site-determining protein XF1321 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug 2000 #text_change 02-Sep-2000
C;Accession: C82695
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717, PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82695
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-269 <SIM>
A;Cross-references: GB:AE003965; GB:AE003849; NID:g9106313; PIDN:AAF84130.1; GSPDB:GN001
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1321
C;Superfamily: cell division inhibitor minD

```
  Query Match             31.1%;  Score 514.5;  DB 2;  Length 269;
  Best Local Similarity   43.5%;  Pred. No. 8.2e-34,
  Matches 120;  Conservative  58;  Mismatches  79;  Indels  19;  Gaps   8;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
Db       4 IIVITSGKGGVGKTTTSASLACGLARRGKKVVVIDFDVGLRNLDLIMGCERRVVYDFVNV 63

Qy     120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
Db      64 IDGEATLKQALIKDKPFDNLYLLAAAQTRDKDALTFEG--VEKVLNELQA--EGF-DYIC 118

Qy     180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLL-------ECDGIRDIKMI 232
Db     119 CDSPAGIEKGASLAMYFADRAVVVVNPEVSSVRDSDRIIGLLDSKTKKAETGGSIITTLL 178

Qy     233 VNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLA 292
Db     179 LTRYSPARVESGEMLSIADVEEVLGLKAIGVIPESGDVLNASNKGEPVILDN-NSLAGLA 237

Qy     293 FEQAAWRLVEQD-SMKAVMVEEEPKKRGFFS-FFGG 326
Db     238 YEDAVGRILGEDHPMRFTTVE----KKGFFSKLFGG 269
```

RESULT 12
C64561
cell division inhibitor minD   Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 02-Feb-2001
C;Accession: C64561
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.,
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O., Fraser, C.M
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: C64561
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-268 <TOM>
A;Cross-references: GB:AE000551; GB:AE000511; NID:g2313436; PIDN:AAD07400.1; PID:g231343
C;Genetics:
A;Gene: minD
C;Function:
A;Description: an ATPase required for correct placement of cell division site
C;Superfamily: cell division inhibitor minD
C;Keywords: ATP; cell division control; nucleotide binding; P-loop
F;10-17/Region: nucleotide-binding motif A (P-loop)
F;16/Binding site: ATP (Lys) #status predicted

```
  Query Match             30.6%;  Score 506.5;  DB 2;  Length 268;
  Best Local Similarity   40.1%;  Pred. No. 3.6e-33;
  Matches 110;  Conservative  65;  Mismatches  80;  Indels  19;  Gaps   5;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
Db       4 VVTITSGKGGVGKSTTTANLAIGLAESGKKVVAVDFDIGLRNLDMILGLENRIVYDVVDV 63

Qy     120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
Db      64 MEKNCNLSQALITDKKTKNLSFLAASQSKDKNILDKEKVAI--LINALR----ADFDYIL 117

Qy     180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIK-------MI 232
Db     118 IDSPAGIESGFEHAILHADMALVVVTPEVSSLRDSDRVVGIIDAKSNRAKKGMEVHKHLI 177

Qy     233 VNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLA 292
Db     178 INRLKPELVANGEMISIEEVLKILCLPLIGIIPEDHHIISATNKGEPVI--RTDCESAKA 235

Qy     293 FEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
Db     236 YQRITRRILGEE----VEYVEFKAKRGFFSALKG 265
```

RESULT 13
B71945
cell division inhibitor - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
C;Accession: B71945
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Jves, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; T
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho

A;Reference number: A71800; MUID:99129557; PMID:9923682
A;Accession: B71945
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-268 <APN>
A;Cross-references: GB:AE001468; GB:AE001439; NID:g4154838; PIDN:AAD05985.1; PID:g415485
A;Experimental source: strain J99
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

Query Match 30.3%; Score 500.5; DB 2; Length 268;
Best Local Similarity 39.8%; Pred. No. 1.1e-32;
Matches 109; Conservative 65; Mismatches 81; Indels 19; Gaps 5;

Qy 60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119
Db 4 VVTITSGKGGVGKSTTTANLAIGLAESGKKVVAVDFDIGLRNLDMILGLENRIVYDVVDV 63

Qy 120 INGDCRLDQALVPDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179
Db 64 MEKNCNLSQALITDKKTKNLSFLAASQSKDKNILDKEKVAI--LINALR - ADFDYIL 117

Qy 180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIR---- - DIKMI 232
Db 118 IDSPAGIESGFEHAILHADMALVVVTPEVSSLRDSDRVIGIIDAKSNRAKSGEEVHKHLI 177

Qy 233 VNPVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLA 292
Db 178 INPLKPELVANGEMISIEEVLKILCLPLIGIIPEDHHIISATNKGEPVI--RTDCESAKA 235

Qy 293 FEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
Db 236 YQPITPPILGEE----VEYVEFKAKRGFFSALKG 265

RESULT 14
AE0253
septum site-determining protein [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C;Accession: AE0253
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0253
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <KUP>
A;Cross-references: GB:AL590842; PIDN:CAC90880.1; PID:g15980088; GSPDB:GN00175
C;Genetics:
A;Gene: minD
C;Superfamily: cell division inhibitor minD

Query Match 29.4%; Score 487; DB 2; Length 270;
Best Local Similarity 41.2%; Pred. No. 1.3e-31;
Matches 114; Conservative 55; Mismatches 90; Indels 18; Gaps 7;

Qy 59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
Db 3 RIIVVTSGKGGVGKTTSSAAIATGLAQKGKKTVVIDFDIGLRNLDLIMGCERRVVYDFVN 62

Qy 119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
Db 63 VIQGDATLNQALIKDKRTDNLYILPASQTRDK-----DALTKEGVEKVLNDLGEMNFEFV 117

Qy 179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIR------DIK-- 230
Db 118 VCDSPAGIESGALMALYFADEAVITTNPEVSSVRDSDPILGILSSKSPRAENSQEPIKEH 177

Qy 231 MIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAG 290
Db 178 LLLTRYNPGRVNKGDMLSMEDVLDILRIPLVGVIPEDQSVLRASNQGEPVILDKESD-AG 236

Qy 291 LAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFS-FFGG 326
Db 237 KAYEDTVDRLLGEE-RPFPFVEEE--KKGFLKPLFGG 270

RESULT 15
AH0724
septum site determining protein [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov 2001
C;Accession: AH0724
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; PMID:11677608
A;Accession: AH0724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05498.1; PID:g16503000; GSPDB:GN00176
C;Genetics:
A;Gene: STY1945
C;Superfamily: cell division inhibitor minD

Query Match 29.2%; Score 483; DB 2; Length 270;
Best Local Similarity 41.2%; Pred. No. 2.8e-31;
Matches 114; Conservative 57; Mismatches 88; Indels 18; Gaps 8;

Qy 59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
Db 3 RIIVVTSGKGGVGKTTSSAAIATGLAQKGKKTVVIDFDIGLRNLDLIMGCERRVVYDFVN 62

Qy 119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
Db 63 VIQGDATLNQALIKDKRTENLFILPASQTRDKDALTREGVAK VLDSLKAM---DFEFI 117

Qy 179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC------DGIPDIK- 230
Db 118 VCDSPAGIETGALMALYFADEAIITTNPEVSSVRDSDPILGILASKSPRAENGEEPIKEY 177

Qy 231 MIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAG 290
Db 178 LLLTRYNPGRVNKGDMLSMEDVLEILRIKLVGVIPEDQSVLRASNQGEPVILDATAD-AG 236

Qy 291 LAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFS FFGG 326
Db 237 KAYADTVDPLLGEE-PPFPFIEEE--KKGFLKPLFGG 270

Search completed: April 16, 2003, 09:08.36
Job time : 46 secs

GenCore version 5.1.4_p5_4578

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:03:53 ; Search time 88 Seconds
(without alignments)
763.311 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLRLFSTNHQSLLLPSSL . . . KAVMVEEEPKKPGFFSFFGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1654	100.0	326	10	Q9MBA2	Q9mba2 arabidopsis
2	1175.5	71.1	295	10	Q9FV40	Q9fv40 tagetes ere
3	1067.5	64.5	306	10	Q9LWY6	Q9lwy6 oryza sativ
4	756	45.7	274	8	Q9T3P6	Q9t3p6 nephroselmi
5	752	45.5	359	8	Q9TJR6	Q9tjr6 prototheca
6	657.5	39.8	268	16	Q8YRJ0	Q8yrj0 anabaena sp
7	606.5	36.7	264	16	Q8RGV1	Q8rgv1 fusobacteri
8	574.5	34.7	264	16	Q8RBB9	Q8rbb9 thermoanaer
9	573.5	34.7	276	16	Q9RWB7	Q9rwb7 deinococcus
10	563	34.0	265	16	Q8XII1	Q8xii1 clostridium
11	559.5	33.8	263	16	Q97JM4	Q97jm4 clostridium
12	546.5	33.0	266	16	Q92BH0	Q92bh0 listeria in
13	545.5	33.0	266	16	Q8Y6Y7	Q8y6y7 listeria mo
14	536.5	32.4	264	16	Q9K8H9	Q9k8h9 bacillus ha
15	516	31.2	262	16	O67033	O67033 aquifex aeo
16	514.5	31.1	269	16	Q9PDQ8	Q9pdq8 xylella fas
17	487	29.4	270	16	Q8ZES6	Q8zes6 yersinia pe
18	483	29.2	270	16	Q8ZP10	Q8zp10 salmonella
19	483	29.2	270	16	Q8Z680	Q8z680 salmonella
20	481.5	29.1	271	16	Q8UAW9	Q8uaw9 agrobacteri
21	480.5	29.1	271	16	Q9HYZ6	Q9hyz6 pseudomonas
22	479	29.0	179	10	Q9SPP0	Q9spp0 oryza sativ
23	468	28.3	276	16	Q9KQN8	Q9kqn8 vibrio chol
24	461.5	27.9	271	16	Q98JG9	Q98jg9 rhizobium l
25	452	27.3	271	16	Q92TZ0	Q92tz0 rhizobium m
26	449.5	27.2	271	2	Q9AG19	Q9ag19 neisseria g
27	449.5	27.2	271	16	Q9JQY6	Q9jqy6 neisseria m
28	440	26.6	271	16	Q8XU29	Q8xu29 ralstonia s
29	388	23.5	245	16	Q8YBH4	Q8ybh4 brucella me
30	380	23.0	264	17	Q8U0H8	Q8u0h8 pyrococcus
31	366	22.1	260	17	Q9V165	Q9v165 pyrococcus
32	344.5	20.8	259	17	O27868	O27868 methanobact
33	341.5	20.6	245	17	O58346	O58346 pyrococcus
34	335	20.3	263	17	O29562	O29562 archaeoglob
35	327.5	19.8	245	17	Q9UYS0	Q9uys0 pyrococcus
36	314.5	19.0	245	17	Q8U3I1	Q8u3i1 pyrococcus
37	307.5	18.6	269	17	Q8TYS4	Q8tys4 methanopyru
38	266	16.1	252	17	O59470	O59470 pyrococcus
39	262.5	15.9	305	17	Q9HNF1	Q9hnf1 halobacteri
40	254	15.4	217	17	O59249	O59249 pyrococcus
41	250	15.1	319	2	Q9L775	Q9l775 xanthomonas
42	249	15.1	280	16	Q9RGE0	Q9rge0 pseudomonas
43	248.5	15.0	253	17	Q8TZX6	Q8tzx6 pyrococcus
44	247.5	15.0	251	17	Q9V1R2	Q9v1r2 pyrococcus
45	246.5	14.9	288	16	Q9KA54	Q9ka54 bacillus ha

ALIGNMENTS

RESULT 1
Q9MBA2
ID Q9MBA2 PRELIMINARY; PRT; 326 AA.
AC Q9MBA2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MinD (Septum site-determining MinD).
GN MIND.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim M., Fujiwara M., Kanamaru K., Tanaka K., Takahashi H.;
RT "Arabidopsis thaliana minD homolog involved in plastid division.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:41-54(1998).
DR EMBL; AB030278, BAA90261.1, -.
DR EMBL; AB009056; BAB08725.1; -.
DR InterPro; IPR000707; ATPase_ParA.
DR Pfam; PF00991; ParA; 1.
SQ SEQUENCE 326 AA; 35690 MW; BE7AEF4307167825 CRC64;

Query Match 100.0%; Score 1654; DB 10; Length 326;
Best Local Similarity 100.0%; Pred. No. 6.4e-128;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60

Qy       61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120

Qy      121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Qy      181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240

Qy      241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Qy      301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
              ||||||||||||||||||||||||||
Db      301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
```

RESULT 2
Q9FV40

```
ID   Q9FV40      PRELIMINARY;      PRT;   295 AA.
AC   Q9FV40;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MinD.
OS   Tagetes erecta (African marigold).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC   Helenieae; Tagetes.
OX   NCBI_TaxID=13708;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Moehs C.P., Tian L., DellaPenna D.;
RT   "Analysis of carotenoid biosynthetic gene expression during marigold
RT   petal development.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF251019; AAG10431.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
SQ   SEQUENCE   295 AA;  32756 MW;  212BD476BF58747F CRC64;

  Query Match             71.1%;  Score 1175.5;  DB 10;  Length 295;
  Best Local Similarity   72.0%;  Pred. No. 1.2e-88;
  Matches  247;  Conservative   15;  Mismatches   16;  Indels   65;  Gaps    7;

Qy        1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVN-NPSRRS-------------PIRSVLQ 46
            | |||                 |.. .|. |  |   ||  ..             ||||||
Db        1 MTSLRF-------------LTEPSLVCSSTFPTFNPLHKTLTKPTPKPYPKPPPIRSVLQ 47

Qy       47 FNRKPELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLL 106
            :||||||||:|||                             |||||||:||||||||
Db       48 YNRKPELAGDTPR-----------------------------VVAIDADVGLRNLDLLL 77

Qy      107 GLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDA 166
            |||||||| |||:||||||||||||||||||||||||||||||||||:||||||| || ||
Db       78 GLENRVNYTVVEVLNGDCRLDQALVRDKRWSNFELLCISKPRSKLPLGFGGKALVWL-DA 136

Qy      167 LKTRPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGI 226
            || | || ||||:||||||||||||||||||||||||||||||||||||||||||||||
Db      137 LKDRQEGCPDFILIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGI 196

Qy      227 RDIKMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDS---EVIRSTNRGFPLVLN 283
            ||||||||||||:|:|||||||||||||||||||        |:   ||||||||||||||
Db      197 RDIKMIVNRVRTDLIPGEDMMSVLDVQEMLGLSLL----SDTPGFEVIRSTNRGFPLVLN 252

Qy      284 KPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 326
            |||||||||||||||||||||||||||||||||||||||||||
Db      253 KPPTLAGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFSFFGG 295
```

RESULT 3
Q9LWY6

```
ID   Q9LWY6      PRELIMINARY;      PRT;   306 AA.
AC   Q9LWY6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Similar to Chlorella vulgaris C-27 chloroplast DNA.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. NIPPONBARE;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT   clone:P0644B06.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP001129; BAA90628.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
SQ   SEQUENCE   306 AA;  32341 MW;  CECAB38219512690 CRC64;

  Query Match             64.5%;  Score 1067.5;  DB 10;  Length 306;
  Best Local Similarity   68.4%;  Pred. No. 9.7e-80;
  Matches  216;  Conservative   31;  Mismatches   50;  Indels   19;  Gaps    5;

Qy       14 LLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRIVVITSGKGGVGKT 73
            |||||         |    ::|:|              ||:|  ||:||:|:|||||||||
Db        7 LLLPS--------RCPPPASSPARHGG-------RTAPELSGPTPRVVVVTSGKGGVGKT 51

Qy       74 TTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRD 133
            ||||.  |||  ||:|||  ||||||||||:|:|||:  |  .|. ||||||||||
Db       52 TTTANLAASLARLSLSAVAVDADAGLRNLDLLLGLENRVHLTAADVLAGDCRLDQALVRH 111

Qy      134 KRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITA 193
            .  : :||.|||||||: || | |  |. ||. |     | ||.|||||.||||:|
Db      112 RALHDLQLLCLSKPRSKLPLAFGSKTLTWVADALR-RAANPPAFILIDCPAGVDAGFVTA 170

Qy      194 ITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDMIKGEDMMSVLDVQ 253
            | || |||||||||||||||||| |||||||:|||:|||||| |::|||||| ||||
Db      171 IAPAEEAVLVTTPDITALRDADRVAGLLECDGIKDIKIIVNRVRPDLVKGEDMMSALDVQ 230

Qy      254 EMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRLVEQDSMKAVMVEE 313
            ||||| ||||:|||:||||||||| ||||| ||| |||| ||| |||||:|:| ||||||
Db      231 EMLGLPLLGVVPEDAEVIRSTNRGVPLVLNDPPTPAGLALEQATWRLVERDAMTAVMVEE 290

Qy      314 E--PKKR-GFFSFFGG 326
            :  |||: ||||||||
Db      291 QERPKKKAGFFSFFGG 306
```

RESULT 4
Q9T3P6

```
ID   Q9T3P6      PRELIMINARY;      PRT;   274 AA.
AC   Q9T3P6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Septum-site determining protein.
GN   MIND.
OS   Nephroselmis olivacea.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;
OC   Chlorodendrales; Chlorodendraceae; Nephroselmis.
```

```
OX   NCBI_TaxID=31312;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99398694; PubMed=10468594;
RA   Turmel M., Otis C., Lemieux C.;
RT   "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT   olivacea: Insights into the architecture of ancestral chloroplast
RT   genomes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Turmel M., Otis C., Lemieux C.;
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF137379; AAD54908.1; -.
DR   EMBL; AF137379; AAD54881.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Chloroplast.
SQ   SEQUENCE   274 AA;  30126 MW;  9D0A1449E6815845 CRC64;

  Query Match             45.7%;  Score 756;  DB 8;  Length 274;
  Best Local Similarity   58.7%;  Pred. No. 3.4e-54;
  Matches  155;  Conservative   39;  Mismatches   60;  Indels   10;  Gaps    3;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           |::|||||||||||||| |||:|: :|| |: |  ||||:|||||||||||||| || :|
Db      14 RVIVITSGKGGVGKTTATANLGMCIARLGYRVALIDADIGLRNLDLLLGLENRVVYTAME 73

Qy     119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
           || | |||:|||:|||||  |  :| :|| | .  |    | .  .||:: | |      :|
Db      74 VIEGQCRLEQALIRDKRWKNLSMLAMSKNRQRYNM- TRKNMMMIVDSIKER---GYQYI 128

Qy     179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRT 238
           :|||||||||||: || ||:||:|||||:|||:|!||| |||| .       .. ||||
Db     129 LIDCPAGIDAGFVNAIAPADEAILVTTPEITAIRDADRVAGLLEANDFYNVRLVANRVRP 188

Qy     239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
           :||:  ||||| ||| |:|: ||| |||| || ||||| ||| | ||||:|||:||
Db     189 EMIQQNDMMSVDDVQGMIGVPLLGAIPEDKNVIISTNRGEPLVCQKTITLAGVAFEEAAR 248

Qy     299 RLVEQDSMKAVMVEEEPKKRGFFS 322
           |||   |       :    ||:|:
Db     249 RLVGLPS-----PSDSAPSRGWFA 267


RESULT 5
Q9TJR6
ID   Q9TJR6      PRELIMINARY;      PRT;   359 AA.
AC   Q9TJR6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Cell division inhibitor MinD.
GN   MIND.
OS   Prototheca wickerhamii.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
OC   Chlorellaceae; Prototheca.
OX   NCBI_TaxID=3111;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=263-11;
RA   Knauf U., Hachtel W.;
RT   "A 22 kb fragment of the 53 kb plastid genome of the colourless alga
RT   Prototheka wickerhamii containing atp-, rpl-,rps-, rrn-, and trn-
RT   genes.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ245645; CAB53105.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Chloroplast.
SQ   SEQUENCE   359 AA;  40804 MW;  E550EAF50BC0A51E CRC64;

  Query Match             45.5%;  Score 752;  DB 8;  Length 359;
  Best Local Similarity   55.4%;  Pred. No. 1e-53;
  Matches  159;  Conservative   44;  Mismatches   62;  Indels   22;  Gaps    5;

Qy      35 PSRRSPIRSVL-QFNRKP-------ELAGE---------TPRIVVITSGKGGVGKTTTTA 77
           |  |  |:  | |  :||        ||  |          |||:|||||||||||||||
Db      53 PEERKIIKEQLEQLIQKPSESEYNTELDIELDKGDSDELEPRVIVITSGKGGVGKTTTTA 112

Qy      78 NVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWS 137
           |:|:|:||:|: |  ||||:|||||||||||||: :| :::| | |||||||||:|||
Db     113 NLGMSIARFGYRVALIDADIGLRNLDLLLGLENRITFTAMDIIEGRCRLDQALVREKRWK 172

Qy     138 NFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDCPAGIDAGFITAITPA 197
           |  || :||  |      : :  || ::|   |   : |:||||||||| ||| || ||
Db     173 NLALLAVSKNHQK--YNVTQQHMRQLVFSIK---ELGINSILIDCPAGIDVGFINAIAPA 227

Qy     198 NEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDMIKGEDMMSVLDVQEMLG 257
           ||::||||:|||:|||||| |||| . | | |···|||| |||·   |:|::|||| ||
Db     228 QEAIIVTTPEITAIRDADRVAGLLEANTIVDTKLLLNRVRMDMIQNSTMLSIMDVQETLG 287

Qy     258 LSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRLVEQD 304
           . |:| ||||. || |||:| ||||.|  ||:|:||| || ||: ::
Db     288 IPLLGAIPEDTNVIISTNKGEPLVLDKKLTLSGIAFENAAPPLIGKE 334


RESULT 6
Q8YRJ0
ID   Q8YRJ0      PRELIMINARY;      PRT;   268 AA.
AC   Q8YRJ0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Septum site-determining protein.
GN   MIND OR ALR3456.
OS   Anabaena sp. (strain PCC 7120).
OC   Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX   NCBI_TaxID=103690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21595285; PubMed=11759840;
RA   Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA   Yasuda M., Tabata S.;
RT   "Complete genomic sequence of the filamentous nitrogen-fixing
RT   cyanobacterium Anabaena sp. strain PCC 7120.";
RL   DNA Res. 8:205-213(2001).
DR   EMBL; AP003593; BAB75155.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   268 AA;  29262 MW;  02C4A9D604C1CBCC CRC64;

  Query Match             39.8%;  Score 657.5;  DB 16;  Length 268;
  Best Local Similarity   54.4%;  Pred. No. 4.1e-46;
  Matches  137;  Conservative   48;  Mismatches   54;  Indels   13;  Gaps    4;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           ||:|||||||||||| :||:|::||: |  |  :||| |||||||||||||: || ||
Db       3 RIIVITSGKGGVGKTTVSANLGMALAKMGRQVALVDADFGLRNLDLLLGLENRIVYTAVE 62

Qy     119 VINGDCRLDQALVRDKRWSNFELLCISKPRSK---LPMGFGGKALEWLVDALKTRPEGSP 175
           |:: :|||:||||:||| |  ||  .. |:|    |      . .: ||:||  : :
Db      63 VLSRECRLEQALVKDKRQPNLVLLPAAQNRTKDAVTP-----EQMKLLVNALAQKYQ--- 114

Qy     176 DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNR 235
            ::||| |||||. ||  || ||:.||||:|:.|||||| ||||  |:: : :|:||
Db     115 -YVIIDSPAGIENGFKNAIAPAKEALIVTTPEISAVRDADPVVGLLEAQGVKPVHLIINP 173

Qy     236 VRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQ 295
           :|  |::  |||||  ||||·| . |:||:|:|  || ||||| ||||   |:|| :| | |
Db     174 IRPAMVRANDMMSVQDVQELLAIPLIGVLPDDERVIVSTNRGEPLVLGDTPSLAAVAVEN 233
```

```
Qy    296 AAWRLVEQDSMK 307
          | || | |···
Db    234 IARRL-EGDTVE 244


RESULT 7
Q8RGV1
ID   Q8RGV1      PRELIMINARY;      PRT;   264 AA.
AC   Q8RGV1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Cell division inhibitor MinD.
GN   FN0176
OS   Fusobacterium nucleatum (subsp. nucleatum).
OC   Bacteria; Fusobacteria; Fusobacterium.
OX   NCBI_TaxID=76856;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25586;
RX   MEDLINE=21886394; PubMed=11889109;
RA   Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA   Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA   Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA   Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA   Fonstein M., Kyrpides N., Overbeek R.;
RT   "Genome sequence and analysis of the oral bacterium Fusobacterium
RT   nucleatum strain ATCC 25586.";
RL   J. Bacteriol. 184:2005 2018(2002).
DR   EMBL; AE010531; AAL94382.1; -.
KW   Complete proteome.
SQ   SEQUENCE   264 AA;  28675 MW;  DA74D05DF8ED1F68 CRC64;

  Query Match             36.7%;  Score 606.5;  DB 16;  Length 264;
  Best Local Similarity   50.0%;  Pred. No. 6.3e-42;
  Matches  121;  Conservative   55;  Mismatches   59,  Indels    7,  Gaps    3,

Qy     59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
          |::|:||||||||||||||:|  ||  |  |. || |·||||||: :|||||: |  |·
Db      4 RVIVVTSGKGGVGKTTTTANIGAGLADKGHKVLLIDTDIGLRNLDVVMGLENRIVYDLVD 63

Qy    119 VINGDCRLDQALVPDKPWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
          ||   ||: || ::|||  |  ||  .. | |  ·    · ·· |·|:|     | |:|
Db     64 VIEERCPISQAFIKDRPCPNLVLLPAAQIRDKNDV--TPEQMKSLIDSLK----ASFDYI 117

Qy    179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRT 238
          ::|||||: |!  |!  |:||' |||·.. | |||||: |||  ||.:: ...||::
Db    118 LVDCPAGIEQGFKNAIVAADEAVVVTTPEVSATRDADRIIGLLEASGIKEPRLVINRLKI 177

Qy    239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
          ||:| · |:|| |    ||: |||| |:|  |: |||·! |||  |  :||  ||:  |
Db    178 DMVKDKNMLSVEDILDILGIKLLGVVPDDETVVISTNKGEPLVY-KGDSLAAKAFKNIAN 236

Qy    299 RL 300
          |:
Db    237 RI 238


RESULT 8
Q8RBB9
ID   Q8RBB9      PRELIMINARY;      PRT;   264 AA.
AC   Q8RBB9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Septum formation inhibitor-activating ATPase.
GN   MIND OR TTE0903.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Thermoanaerobacteriales; Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4T / JCM11007;
RX   MEDLINE=21992816, PubMed=11997336,
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
DR   EMBL; AE013057; AAM24159.1; -.
KW   Complete proteome.
SQ   SEQUENCE   264 AA;  29357 MW;  3F85B51D77C6859C CRC64;

  Query Match             34.7%;  Score 574.5;  DB 16,  Length 264,
  Best Local Similarity   46.3%;  Pred. No. 2.7e-39;
  Matches  113;  Conservative   56,  Mismatches   68,  Indels    7,  Gaps    3,

Qy     61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
          :|||||||||||||||:|  ||  |: |  :| |:||||||:::|||||: |  |:|:
Db      5 IVITSGKGGVGKTTTTANIGTYLAMKGYKVALVDTDIGLRNLDVVMGLENRIVYDIVDVV 64

Qy    121 NGDCRLDQALVPDKPWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
          | ||| |||:|||:    ||  :  | |          |   |:       | | ·
Db     65 EGQCPLKQALIPDKPFDTLYLLPAAQTRDKTAV--TPEQMRKLIQDLKE----EFDYILV 118

Qy    181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVPTDM 240
          |||||  ||  :|  |  |·:|||·· :|:|||: :|||     .  :..:|.. ||
Db    119 DCPAGIEQGFKNAIAGADRAIVVTTPEVSAVRDADRIIGLLEAAELHNPLLVINRIKMDM 178

Qy    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
          :|  ||| : |: : :  . |||||:|  :| ||| | : :|     ||  |:     |
Db    179 VKRGDMMDIEDIIDILAIDLLGVIPDDENIIISTNKGEPIVLDE-KSLASQAYRNLVERF 237

Qy    301 VEQD 304
          ::::
Db    238 LDRN 241


RESULT 9
Q9RWB7
ID   Q9RWB7      PRELIMINARY;      PRT;   276 AA.
AC   Q9RWB7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Septum site-determining protein.
GN   DR0752.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266,
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001931; AAF10331.1; -.
DR   TIGR; DR0752; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   InterPro; IPR000392; NitrogenaseII.
DR   Pfam; PF00142; fer4_NifH; 1.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   276 AA;  29420 MW;  AD74FDCF45820D0C CRC64,
```

```
  Query Match           34.7%;  Score 573.5;  DB 16;  Length 276;
  Best Local Similarity  44.3%;  Pred. No. 3.5e-39;
  Matches  125;  Conservative   52;  Mismatches   72;  Indels   33;  Gaps    5;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
           :::|·|||||||||||||||·| ·||| |  || || | ·|||·|| : |||.|| :  |· 
Db      12 KVIVVTSGKGGVGKTTTTANIGAALARLGEKVVVIDVDVGLRNLDVVMGLESRVVFDLVD 71

Qy     119 VINGDCRLDQALVRDKRWSNPELLCISKPRSKLPMGFGGKALEWLVDALKTR-------- 170
           |: | ||::|||·|||| |  ||  |: | |              |||
Db      72 VLEGKCRMNQALIRDKRVENLHLLPASQTRDK-------------DALDPEVFKEVVKG 117

Qy     171 --PEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRD 228
             |   | ::|| ||||::|| ||  ||  |::|  |::::|||||: |||   | :
Db     118 LLEEEGFDRVLIDSPAGIESGFRTAAAPAEGALVVVNPEVSSVRDADRIIGLLEAQQITE 177

Qy     229 IKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTL 288
           |:::|||·|  |.   |·|· |.  ·:||·  :|::|||  :| ||| | | || |  |
Db     178 IRLVVNRLRPKMVASGNMLSIDDMVDILGVKPIGIVPEDEGIIVSTNVGEPAVLGK--TK 235

Qy     289 AGLAFEQAAWRLVEQDSMKAVMVEEEPKKRGFFS----FFGG 326
           || ||   | |:  ||     · |||   :| ·:     |||
Db     236 AGDAFMATAQRIQGQDVPFPKKLTEEE---KGIWAAIRRLFGG 274


RESULT 10
Q8XII1
ID   Q8XII1      PRELIMINARY;      PRT;   265 AA.
AC   Q8XII1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Septum site-determining protein.
GN   MIND OR CPE2138.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / TYPE A;
RX   PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; AP003193; BAB81844.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   265 AA;  29086 MW;  23C96B312D068DDE CRC64;

  Query Match           34.0%;  Score 563;  DB 16;  Length 265;
  Best Local Similarity  45.0%;  Pred. No. 2.4e-38;
  Matches  118;  Conservative   55;  Mismatches   79;  Indels   10;  Gaps    4;

Qy      61 VVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           :|||||||||||||||||:| :||  |  || :| | ||||||:|:|||||: || ::||
Db       5 IVITSGKGGVGKTTTTANIGTALAAQGKKVVVVDGDTGLRNLDVLMGLENRIVYTVIDVI 64

Qy     121 NGDCRLDQALVRDKRWSNPELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
             ||  |||·|||·|·|  ||  ·· ·|          ·  ·|· ·|       |· |:
Db      65 ENRCRTKQALIRDKRFNNLYLLPTAQTKDK--NDISPEQMLRLVNELKE----EFDYVIL 118

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVPTDM 240
           ||||||: ||  ||  |: |::|  ·|·||·:|||||  | |:  |· · ·:|||·  ·|
Db     119 DCPAGIEQGFENAIVGADRAIVVVNPEITSVRDADRVIGKLDAKGLENHEVIVNRLNYEM 178

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
             |  ||: : |: | | : |||·|·|  ·  |||·| |·||··  : || ||     |:
Db     179 TKKGDMLDISDIIETLSVKLLGVVPDDRNITVSTNKGEPIVLDEKAS-AGQAFRNIGRRI 237

Qy     301 VEQDSMKAVMVEEEPKKRGFFS 322
           : :|     :::   : :| ||
Db     238 IGED---VPIMDLNTEHQGIFS 256


RESULT 11
Q97JM4
ID   Q97JM4      PRELIMINARY;      PRT;   263 AA.
AC   Q97JM4;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Septum site-determining protein MinD, ATPase.
GN   CAC1249.
OS   Clostridium acetobutylicum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC   Clostridiales; Clostridiaceae; Clostridium.
OX   NCBI_TaxID=1488;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX   MEDLINE=21359325; PubMed=11466286;
RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA   Bennett G.N., Koonin E.V., Smith D.R.;
RT   "Genome sequence and comparative analysis of the solvent-producing
RT   bacterium Clostridium acetobutylicum.";
RL   J. Bacteriol. 183:4823-4838(2001).
DR   EMBL; AE007638; AAK79221.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   263 AA;  28769 MW;  B06AFBEA9DCF3D73 CRC64;

  Query Match           33.8%;  Score 559.5;  DB 16;  Length 263;
  Best Local Similarity  44.7%;  Pred. No. 4.6e-38;
  Matches  117;  Conservative   57;  Mismatches   77;  Indels   11;  Gaps    4;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           :|||||||||||||||||:| :||     || :| | ||||||:|:|||||: :| ::||·
Db       5 IVITSGKGGVGKTTTTANIGTALAAMKKKVVLVDGDTGLRNLDVLMGLENRIVFTLLDVV 64

Qy     121 NGDCRLDQALVRDKRWSNPELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
             |:||| |||::|| : |  ||  :: | |  :    : :  ||: ||       |::||
Db      65 EGNCRLKQALIKDKHYENLALLPTAQTRDKNDV--KPEQMLKLVNELKE----EFDYVII 118

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVPTDM 240
           ||||||: ||  |:  |: |::|  |::|::||||| | ::  || | ::||||:  :|
Db     119 DCPAGIEQGFENAVIGADRAIVVVNPEVTSVRDADRVIGKIDAKGIEDHQVIVNRIDYEM 178

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
           :|  ||: : || : | : |:||:| | ::  |||:| |:|||:    || ||    | |:
Db     179 VKRGDMLGIEDVIDNLAIKLIGVVPNDKQITVSTNKGEPIVLNQNAN-AGKAFRDIARRV 237

Qy     301 VEQDSMKAVMVEEEPKKRGFFS 322
           . ..     |  |:   . || :
Db     238 LGEE----VPFEKYETQTGFIA 255


RESULT 12
Q92BH0
ID   Q92BH0      PRELIMINARY;      PRT;   266 AA.
AC   Q92BH0;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   MinD protein.
GN   MIND OR LIN1579.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
```

```
RN   [1]
RP   SEQUENCE FROM N A
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F , Berche P , Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L -M , Kaerst U , Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A , Mata Vicente J , Ng E., Nedjari H.,
RA   Nordsiek G., Novella S , de Pablos B , Perez-Diaz J.-C., Purcell R.,
RA   Remmel B , Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596169, CAC96810.1, -.
DR   ListiList; LIN01579; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   266 AA;  29074 MW;  D45807D1E02B4E97 CRC64,

  Query Match             33.0%;  Score 546.5;  DB 16;  Length 266;
  Best Local Similarity   43.3%;  Pred. No. 5.5e-37;
  Matches  117;  Conservative   60;  Mismatches   80;  Indels   13;  Gaps    6;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
            :||||||||||||:|||:| :|| |  |  || |:|||||::||||||: |  |:|:
Db       5 IVITSGKGGVGKTTSTANLGTALALQGKKVCLIDMDIGLRNLDVVLGLENRIIYDLVDVV 64

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
             | |.. ||: ||| .  |      ::        |. :   |: :  |  ||:   |:|:|
Db      65 EGPCKIHQAMIKDKRFDDL-LFLLPAAQTTDKNAVSGEQMIELINQL--RPD--YDYILI 119

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIPDIKMIVNRVRTDM 240
           ||||||: |.  |    |:  |..|||:|:|:|||||:  |        |:|:|:|:|| |
Db     120 DCPAGIETGYKNAVAGADKAIVVTTPEISAVRDADRIIGLLEKEDIEPPKLIINRIRTQM 179

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
            :    |:|: ::    | .  ||:|| :|||||:| | |: :  |     :    | |:
Db     180 MVNGDVMDIDEITTHLSIELLGIIIDDDEVIPSSNSGDPVAM-LPNNRASQGYRNIARRI 238

Qy     301 VEQDSMKAVMVEEEPKKRGFFS----FFGG 326
            : :      :: | || |||:      |||
Db     239 LGE---SIPLMSIEAKKPGFFTRLKQLFGG 265
```

```
RESULT 13
Q8Y6Y7
ID   Q8Y6Y7      PRELIMINARY;      PRT;   266 AA.
AC   Q8Y6Y7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   MinD protein.
GN   MIND OR LMO1544.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N A.
RC   STRAIN=EGD-E / SEROVAR 1/2A;
RX   MEDLINE=21537279, PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E , Dominguez-Bernal G , Duchaud E , Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U , Kreft J , Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294.849-852(2001).
DR   EMBL; AL591979; CAC99622.1; -.
DR   ListiList; LMO01544; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   266 AA;  29066 MW;  0CEE19E77A7F278A CRC64,

  Query Match             33.0%;  Score 545.5,  DB 16,  Length 266;
  Best Local Similarity   43.9%;  Pred. No. 6.6e-37;
  Matches  115;  Conservative   59;  Mismatches   79;  Indels    9;  Gaps    5;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
            :|||||||||||||:|||:| :||  |  |  || |:|||||::||||||: |  |:|:
Db       5 IVITSGKGGVGKTTSTANLGTALALQGKKVCLIDMDIGLRNLDVVLGLENRIIYDLVDVV 64

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
             | |:: || : |||: :  |                |:  |.: |  ||:   |:| |
Db      65 EGPCKIHQAMIKDKRFDDL-LFLLPAAQTTDKNAVSGEQMVDLINQL--RPD  YDFILI 119

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
           ||||||: |:  |:  |..|::||||.|:|||||:  |||    |  | |.||.|| |
Db     120 DCPAGIETGYKNAVAGADKAIVVTTPEISAVRDADRIIGLLEKEDIEPPKLIINRIRTQM 179

Qy     241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
            :   |:| : :.   | . ||:|. :| |||||·| | |: ·  |  |   ·   | |:
Db     180 MMNGDVMDIDEITTHLSIELLGIIIDDDEVIPSSNSGDPVAM-LPNNRASQGYPNIARRI 238

Qy     301 VEQDSMKAVMVEEEPKKRGFFS 322
            : .       .: | || |||:
Db     239 LGE---SIPLMSIETKKAGFFA 257
```

```
RESULT 14
Q9K8H9
ID   Q9K8H9      PRELIMINARY;      PRT,   264 AA.
AC   Q9K8H9;
DT   01 OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Septum site-determining protein.
GN   MIND OR BH3027.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582, PubMed=11058132,
RA   Takami H , Nakasone K , Takaki Y , Maeno G , Sasaki R , Masui N ,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; AP001517; BAB06746.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam, PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   264 AA;  28818 MW;  78D523EBE583172F CRC64,

  Query Match             32.4%;  Score 536.5,  DB 16,  Length 264;
  Best Local Similarity   41.1%;  Pred. No. 3.6e-36;
  Matches  109;  Conservative   66,  Mismatches   83,  Indels    7;  Gaps    3;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
            :||||||||||||||:||:| :|| |  |  | | |||||| . ||||· | | |·
```

```
Db      5 IVITSGKGGVGKTTTSANIGTALALSGKKVCLVDTDIGLRNLDVVMGLENRIIYDLVDVV 64

Qy    121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Db     65 EGRCRLKQALIKDKRFECLNLLPAAQTKDK--SAVTPEQMKEIVEELKQ----EYDYVLI 118

Qy    181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240

Db    119 DCPAGIEQGFKNAVAGADKAIVVTTPEISSVRDADRIIGLLEKEEVEAPRLVVNRIRGHM 178

Qy    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Db    179 MKNGEMLDVDEIVSILAIELLGIVVDDENVIKFSNKGEPIALH-PDSKASVAYRNIARRI 237

Qy    301 VEQDSMKAVMVEEEPKKRGFFSFFG 325

Db    238 LGETVPLMSFEQEKGVLAKIKSFFG 262


RESULT 15
O67033
ID   O67033      PRELIMINARY;      PRT;   262 AA.
AC   O67033;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Septum site-determining protein MIND.
GN   MIND2 OR AQ_877.
OS   Aquifex aeolicus.
OC   Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX   NCBI_TaxID=63363;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VF5;
RX   MEDLINE=98196666; PubMed=9537320;
RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT   "The complete genome of the hyperthermophilic bacterium Aquifex
RT   aeolicus.";
RL   Nature 392:353-358(1998).
DR   EMBL; AE000712; AAC06996.1; -.
DR   InterPro; IPR000707; ATPase_ParA.
DR   Pfam; PF00991; ParA; 1.
KW   Complete proteome.
SQ   SEQUENCE   262 AA;  28913 MW;  535C2E5F3D5B352A CRC64;

  Query Match             31.2%;  Score 516;  DB 16;  Length 262;
  Best Local Similarity   41.4%;  Pred. No. 1.7e-34;
  Matches  108;  Conservative   65;  Mismatches   70;  Indels   18;  Gaps    3;

Qy     60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEV 119

Db      4 VIVITSGKGGVGKTTLTANIGTALAKLGKKVLLIDADIGLRNLDMILGLENRIVYDILDV 63

Qy    120 INGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFII 179

Db     64 LEGRVPYEKALVKDKRGLSLWLLPANQRANKDVIDIE----KWNKTVEEIKNSGNYDYIL 119

Qy    180 IDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTD 239

Db    120 VDSPAGIEKGFQIAVSPADKALIVVNPEVSSIRDADRVIGLLESMDKRNYKVIVNRIKWE 179

Qy    240 MIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNK--PPTLAGLAFEQAA 297

Db    180 MVKRGAMLSVEDIVDILKAEIIGIIPEEPKLVDFTNRGEPIVLDEKFPASQAII------ 233

Qy    298 WRLVEQDSMKAVMVEEEPKKR 318

Db    234 ------DTARRLMGESIPLKR 248


Search completed: April 16, 2003, 09:07:46

Job time : 91 secs
```

OM protein - protein search, using sw model

Run on: April 16, 2003, 09:05:27 ; Search time 20 Seconds
(without alignments)
1232.506 Million cell updates/sec

Title: US-10-067-989-1
Perfect score: 1654
Sequence: 1 MASLRLFSTNHQSLLLPSSL.....KAVMVEEEPKKRGFFSFFGG 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1654	100.0	326	12	US-10-067-989-1	Sequence 1, Appli
2	792	47.9	284	12	US-10-067-989-2	Sequence 2, Appli
3	643	38.9	266	12	US-10-067-989-3	Sequence 3, Appli
4	479	29.0	270	12	US-10-067-989-4	Sequence 4, Appli
5	197.5	11.9	279	9	US-09-738-626-6925	Sequence 6925, Ap
6	190	11.5	390	9	US-09-712-363-192	Sequence 192, App
7	176	10.6	288	9	US-10-102-806-730	Sequence 730, App
8	173.5	10.5	375	9	US-09-738-626-4744	Sequence 4744, Ap
9	164.5	9.9	478	9	US-09-738-626-3883	Sequence 3883, Ap
10	145.5	8.8	253	10	US-09-815-242-10915	Sequence 10915, A
11	134	8.1	264	10	US-09-815-242-11559	Sequence 11559, A
12	128.5	7.8	225	10	US-09-767-041-14	Sequence 14, Appl
13	118	7.1	261	9	US-09-738-626-5070	Sequence 5070, Ap
14	107.5	6.5	308	10	US-09-791-171-94	Sequence 94, Appl
15	107.5	6.5	666	10	US-09-791-171-70	Sequence 70, Appl
16	102	6.2	261	10	US-09-841-132-517	Sequence 517, App
17	101.5	6.1	340	9	US-09-738-626-5233	Sequence 5233, Ap
18	99.5	6.0	262	10	US-09-815-242-5087	Sequence 5087, Ap
19	94	5.7	6304	9	US-10-147-026-16	Sequence 16, Appl
20	91	5.5	644	9	US-09-948-137A-2	Sequence 2, Appli
21	91	5.5	644	10	US-09-815-242-10318	Sequence 10318, A
22	90.5	5.5	335	9	US-09-738-626-6510	Sequence 6510, Ap
23	90.5	5.5	341	9	US-09-738-626-3835	Sequence 3835, Ap
24	89	5.4	453	10	US-09-815-242-10263	Sequence 10263, A
25	88	5.3	612	10	US-09-925-301-1358	Sequence 1358, Ap
26	87	5.3	421	10	US-09-815-242-11192	Sequence 11192, A
27	86.5	5.2	416	9	US-09-860-846-39	Sequence 39, Appl
28	86.5	5.2	416	10	US-09-861-289-39	Sequence 39, Appl
29	84	5.1	297	9	US-10-098-807-5	Sequence 5, Appli
30	83.5	5.0	996	9	US-09-738-626-3847	Sequence 3847, Ap
31	83.5	5.0	3354	9	US-10-160-758-11	Sequence 11, Appl
32	83.5	5.0	3354	9	US-10-160-758-12	Sequence 12, Appl
33	83	5.0	229	10	US-09-767-041-13	Sequence 13, Appl
34	83	5.0	397	9	US-09-712-363-280	Sequence 280, App
35	83	5.0	501	10	US-09-815-242-14089	Sequence 14089, A
36	83	5.0	577	9	US-09-738-626-4578	Sequence 4578, Ap
37	83	5.0	644	10	US-09-815-242-13760	Sequence 13760, A
38	83	5.0	1256	9	US-10-047-542-90	Sequence 90, Appl
39	82.5	5.0	1080	10	US-09-781-558-4	Sequence 4, Appli
40	82.5	5.0	1115	9	US-10-106-534-7	Sequence 7, Appli
41	82.5	5.0	1135	10	US-09-737-149-35	Sequence 35, Appl
42	82	5.0	258	10	US-09-954-314-12	Sequence 12, Appl
43	82	5.0	369	9	US-09-712-363-170	Sequence 170, App
44	82	5.0	457	10	US-09-815-242-11948	Sequence 11948, A
45	82	5.0	473	9	US-09-738-626-3630	Sequence 3630, Ap

ALIGNMENTS

```
RESULT 1
US-10-067-989-1
; Sequence 1, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
;  APPLICANT: Dinkins, Randy
;  APPLICANT:  Reddy, M.S. Srinivasa
;  APPLICANT:  Collins, Glenn B.
;  TITLE OF INVENTION: Transgenic plants expressing MinD or MinE and an efficient
;  TITLE OF INVENTION:  method for plant chloroplast transformation and gene expression
;  FILE REFERENCE: 028750-219
;  CURRENT APPLICATION NUMBER: US/10/067,989
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/267,488
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 326
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-10-067-989-1

  Query Match           100.0%;  Score 1654;  DB 12;  Length 326;
  Best Local Similarity  100.0%;  Pred. No. 7.8e-152;
  Matches  326;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MASLRLFSTNHQSLLLPSSLSQKTLISSPRFVNNPSRRSPIRSVLQFNRKPELAGETPRI 60

Qy      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVI 120

Qy     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFIII 180

Qy     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 DCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDM 240
```

```
Qy    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 IKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300

Qy    301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326
          |||||||||||||||||||||||||||
Db    301 VEQDSMKAVMVEEEPKKRGFFSFFGG 326


RESULT 2
US-10-067-989-2
; Sequence 2, Application US/10067989
; Patent No  US20020144309A1
; GENERAL INFORMATION:
;  APPLICANT: Dinkins, Randy
;  APPLICANT:  Reddy, M.S. Srinivasa
;  APPLICANT:  Collins, Glenn B.
;  TITLE OF INVENTION: Transgenic plants expressing MinD or MinE and an efficient
;  TITLE OF INVENTION:  method for plant chloroplast transformation and gene expression
;  FILE REFERENCE: 028750-219
;  CURRENT APPLICATION NUMBER: US/10/067,989
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/267,488
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Chlorella vulgaris
US-10-067-989-2

  Query Match             47.9%;  Score 792;  DB 12;  Length 284;
  Best Local Similarity   60.4%;  Pred. No. 1.3e-68;
  Matches  165;  Conservative   39;  Mismatches   55;  Indels   14;  Gaps    5;

Qy     59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
          |:.|||||||||||||||| | || |  |  |||| |||||||||||||| || :.
Db     19 RVIVITSGKGGVGKTTTTANLGMSIARLGYRVALIDADIGLRNLDLLLGLENRVLYTAMD 78

Qy    119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
          :: | |||||||:|||| |  || ||| | |       | :: |:|::|   |     |:
Db     79 IVEGQCRLDQALIPDKRWKNLALLAISKNPQK--YNVTPKNMQNLIDSVK---ELGFQFV 133

Qy    179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIPDIKMIVNRVPT 238
          :|||||||.||| || | |||:|||:|||||| ||||  .|  | .||||
Db    134 LIDCPAGIDVGFINAIASAQEAVIVTTPEITAIRDADRVAGLLEANGIYNVKLLVNRVRP 193

Qy    239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
          |||:  ||||| |||||||: ||| |||: || |||:| ||||||  ||:|:|||  ||
Db    194 DMIQKNDMMSVRDVQEMLGIPLLGAIPEDTSVIISTNKGEPLVLNKKLTLSGIAFENAAR 253

Qy    299 RLV-EQDSMKAVMVEEEPKKRGFF----SFFGG 326
          ||: :||      ::    ::| |      || |
Db    254 RLIGKQD----YFIDLTSPQKGMFQKLQEFFLG 282


RESULT 3
US-10-067-989-3
; Sequence 3, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
;  APPLICANT: Dinkins, Randy
;  APPLICANT:  Reddy, M.S. Srinivasa
;  APPLICANT:  Collins, Glenn B.
;  TITLE OF INVENTION: Transgenic plants expressing MinD or MinE and an efficient
;  TITLE OF INVENTION:  method for plant chloroplast transformation and gene expression
;  FILE REFERENCE: 028750-219
;  CURRENT APPLICATION NUMBER: US/10/067,989
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/267,488
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 266
;   TYPE: PRT
;   ORGANISM: Synnechocystis
US-10-067-989-3

  Query Match             38.9%;  Score 643;  DB 12;  Length 266;
  Best Local Similarity   52.8%;  Pred. No. 3e-54;
  Matches  130;  Conservative   51;  Mismatches   59;  Indels    6;  Gaps    2;

Qy     59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
          ||:|:|||||||||||||.| :||| |  || |||| |||||||||||| |: || ::
Db      3 RIIVVTSGKGGVGKTTTTANLGAALARLGKKVVLIDADFGLRNLDLLLGLEQKIVYTAID 62

Qy    119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
          |:  :| :|:|||:|||  |  ||  :: |||       : :: ||: || :     |:|
Db     63 VLADECTIDKALVKDKRLPNLVLLPAAQNRSK--DAINAEQMQSLVEQLKDK----FDYI 116

Qy    179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRT 238
          |||||||.|||  |. || ||:.|||...|:|||||| |||| . |  | .||||:|
Db    117 IIDCPAGIEAGFRNAVAPAQEAIIVTTPEMSAVRDADRVIGLLEAEDIGKISLIVNRLRP 176

Qy    239 DMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAW 298
          :|::   |:|| |: ·· | · |:|::|:| ·· | |||·| |||: :  :: ||||:  |
Db    177 EMVQLNQMISVEDILDLLAVPLIGILPDDQKIIISTNKGEPLVMEEKLSVPGLAFQNIAR 236

Qy    299 RLVEQD 304
          ||  ||
Db    237 RLEGQD 242


RESULT 4
US-10-067-989-4
; Sequence 4, Application US/10067989
; Patent No. US20020144309A1
; GENERAL INFORMATION:
;  APPLICANT: Dinkins, Randy
;  APPLICANT:  Reddy, M.S. Srinivasa
;  APPLICANT:  Collins, Glenn B.
;  TITLE OF INVENTION: Transgenic plants expressing MinD or MinE and an efficient
;  TITLE OF INVENTION:  method for plant chloroplast transformation and gene expression
;  FILE REFERENCE: 028750-219
;  CURRENT APPLICATION NUMBER: US/10/067,989
;  CURRENT FILING DATE:  2002-05-24
;  PRIOR APPLICATION NUMBER: US 60/267,488
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 270
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-067-989-4

  Query Match             29.0%;  Score 479;  DB 12;  Length 270;
  Best Local Similarity   39.9%;  Pred. No. 2.1e-38;
  Matches  113;  Conservative   58;  Mismatches   82;  Indels   30;  Gaps    8;

Qy     59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNYTCVE 118
          ||:|:|||||||||||::| :   ||: |   | || |:|||||::| | || |  |
Db      3 RIIVVTSGKGGVGKTTSSAAIATGLAQKGKKTVVIDFDIGLRNLDLIMGCEPRVVYDFVN 62

Qy    119 VINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEWLVDALKTRPEGSPDFI 178
          || ||  |:|||::|||  |  :|  |: | |  :   | |   ::| ||       :||
Db     63 VIQGDATLNQALIKDKRTENLYILPASQTRDKDALTREGVAK--VLDDLKAM---DFEFI 117

Qy    179 IIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLEC------DGIRDIK-- 230
          : | ||||: | : |:  | ||:: | |:.. || ||: | |        .|   ||
Db    118 VCDSPAGIETGALMALYFADEAIITTNPEVSSVRDSDRILGILASKSRRAENGEEPIKEH 177

Qy    231 MIVNRVRTDMIKGEDMMSVLDVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAG 290
```

```
Db    178 LLLTRYNPGPVSPGDMLSMEDVLEILPIKLVGVIPEDQSVLRASNQGEPVILDINAD-AG 236

Qy    291 LAFEQAAWRLVEQDSMKAVMVEEEP------KKRGFFS-FFGG 326
Db    237 KAY---------ADTVEPLLGEEPPFRRPIEEEKKGFLKRLFGG 270


RESULT 5
US-09-738-626-6925
; Sequence 6925, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6925
;   LENGTH: 279
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-6925

  Query Match            11.9%;  Score 197.5;  DB 9;  Length 279;
  Best Local Similarity  26.4%;  Pred. No. 3.6e-11;
  Matches   71;  Conservative   46;  Mismatches  109;  Indels   43;  Gaps   10;

Qy     58 PRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRV-NYTC 116
Db      7 PRLITIANQKGGVGKTTSTVNLAASLAIHGLKVLVVDLD-PQGNASTALGVEHRSGTLSS 65

Qy    117 VEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMGFGGKALEW---------LVDAL 167
Db     66 YELLIGECTADEAM--QPSTANENLFCI-----PATLDLAGAEIELVSLVRREYRLADAL 118

Qy    168 KTR--PEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRDADRVTGLLECDG 225
Db    119 GREFIDKHDFDYMIIDCPPSLGLLTINAMTAVNEVLIPIQCEYYALEGVGQL-----LNN 173

Qy    226 IRDIKMIVNR-------------VRTDMIKGEDMMSVLDVQEMLGLSLLG-VIPEDSEVI 271
Db    174 ITMLRQHLNRQLHISAILLTMYDARTNLAE----QVATEVNDHFGDVVLGNKIPRSVKVS 229

Qy    272 RSTNRGFPLVLNKPPTLAGLAFEQAAWRL 300
Db    230 EAPGYGQTVIEYDPGSRGAMAYLDAAKEL 258


RESULT 6
US-09-712-363-192
; Sequence 192, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
;  APPLICANT: Eisenberg, David
;  APPLICANT:  Rotstein, Sergio H.
;  APPLICANT:  Marcotte, Edward M.
;  TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
;  TITLE OF INVENTION:  INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
;  FILE REFERENCE: 07419-032001
;  CURRENT APPLICATION NUMBER: US/09/712,363
;  CURRENT FILING DATE:  2000-11-13
;  PRIOR APPLICATION NUMBER: PCT/US00/02246
;  PRIOR FILING DATE: 2000-01-28
;  PRIOR APPLICATION NUMBER: 60/179,531
;  PRIOR FILING DATE: 2000-02-01
;  PRIOR APPLICATION NUMBER: 60/117,844
;  PRIOR FILING DATE: 1999-01-29
;  PRIOR APPLICATION NUMBER: 60/118,206,
;  PRIOR FILING DATE. 1999-02-01
;  PRIOR APPLICATION NUMBER: 60/126,593
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/134,093
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/134,092
;  PRIOR FILING DATE: 1999-05-14
;  PRIOR APPLICATION NUMBER: 60/165,124
;  PRIOR FILING DATE: 1999-11-12
;  PRIOR APPLICATION NUMBER: 60/165,086
;  PRIOR FILING DATE: 1999-11-12
;  NUMBER OF SEQ ID NOS: 292
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 192
;   LENGTH: 390
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-712-363-192

  Query Match            11.5%;  Score 190;  DB 9;  Length 390;
  Best Local Similarity  24.0%;  Pred. No. 3e-10;
  Matches   68;  Conservative   58;  Mismatches  117;  Indels   40;  Gaps    9;

Qy     32 VNNPSRRSPIRSVLQFN-RKPELAGETP----RIVVITSGKGGVGKTTTTANVGLSLARY 86
Db     95 VMSDEQRTELRKQLRGDTREPVIPFAQPDSLTRVYAVASGKGGVGKSTVTVNLAAAMAVR 154

Qy     87 GFSVVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISK 146
Db    155 GLSIGVLDADIHGHSIPRMMGTTDRPTQV-----------ESMILPPIAHQVKVISIAQ 202

Qy    147 -PRSKLPMGFGGKAL-----EWLVDALKTRPEGSPDFIIIDCPAGID--AGFITAITPAN 198
Db    203 FTQGNTPVVWRGPMLHRALQQFLADVY----WGDLDVLLLDLPPGTGDVAISVAQLIPNA 258

Qy    199 EAVLVTTPDITALRDADRVTGLLECDGIRDIKMIVNRVRTDMIKGEDMMSVLD-----VQ 253
Db    259 ELLVVTTPQLAAAEVAERAGSIALQTRQRIVGVVENMSGLTLPDGTTMQVFGEGGGRLVA 318

Qy    254 EMLG------LSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAG 290
Db    319 ERLSRAVGADVPLLGQIPLDPALVAAGDSGVPLVLSSPDSAIG 361


RESULT 7
US-10-102-806-730
; Sequence 730, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 730
;   LENGTH: 288
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-102-806-730

  Query Match             10.6%;  Score 176;  DB 9;  Length 288;
  Best Local Similarity   22.8%;  Pred. No. 4.5e-09;
  Matches   68;  Conservative   53;  Mismatches  135;  Indels   42;  Gaps    8;

Qy       20 FRFVNNPSRRSPIFSVLQFNRKP-ELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYG 87
Db        2 PAGVTELQPRAPGGGGMEAAAEPGNLAG-VRHIILVLSGKGGVGKSTISTELALALRHAG 60

Qy       88 FSVVAIDADLGLRNLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKP 147
Db       61 KKVGILDVDLCGPSIPRMLGAQGRAVHQC---------------DRGWAPV-FLDREQS 103

Qy      148 RSKLPMGF------------GGKALEWLVDALKTRPEGSPDFIIIDCPAGID---AGFIT 192
Db      104 ISLMSVGFLLEKPDEAVVWRGPKKNALIKQFVSDVAWGELDYLVVDTPPGTSDEHMATIE 163

Qy      193 AITPANE--AVLVTTPDITALRDADRVTGLLECDGIPDIKMIVNPVRTDMIKGEDMMSVL 250
Db      164 ALRPYQPLGALVVTTPQAVSVGDVRRELTFCRKTGLRVMGIVENMSGFTCPHCTECTSVF 223

Qy      251 ------DVQEMLGLSLLGVIPEDSEVIRSTNRGFPLVLNKPPTLAGLAFEQAAWRLVE 302
Db      224 SRGGGEELAQLAGVPFLGSVPLDPALMRTLEEGHDFIQEFPGSPAFAALTSIAQKILD 281


RESULT 8
US-09-738-626-4744
; Sequence 4744, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4744
;   LENGTH: 375
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-4744

  Query Match             10.5%;  Score 173.5;  DB 9;  Length 375;
  Best Local Similarity   24.3%;  Pred. No. 1.1e-08;
  Matches   66;  Conservative   53;  Mismatches   88;  Indels   65;  Gaps   13;

Qy       55 GETPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRNLDLLLGLENRVNY 114
Db      110 GSTTRVYAVASGKGGVGKSSMTVNLAAALAKRGLSVGILDADIYGHSVPGMLGSDQRPH- 168

Qy      115 TCVEVINGDCRLDQALVRDKRWSNFELLCISK-PRSKLPMGFGGKAL-----EWLVDALK 168
Db      169 ---------QVDD-MIMPPQAHGVKMISIAHFTEGNAPVVWRGPMLHRAIQQFLTDVF- 216

Qy      169 TRPEGSPDFIIIDCPAGIDAGFITA--ITPANEAVLVTTPDITALRDADR      216
Db      217 ---WGDLDILLLDLPPGTGDIAITVAQLIPNAELLIVTTPQAAAAEVAERAGTISVQTNQ 273

Qy      217 -VTGLLE-------CDG-IRDI------KMIVNRVRTDMIKGEDMMSVLDVQEMLGLSLL 261
Db      274 KVAGVIENMSAMVLPDGTTMDVFGTGGGQKIADRLTA--VTGEE-----------VKVI 319

Qy      262 GVIPEDSEVIRSTNRGFPLVLNKP--PTLAGL 291
Db      320 GSVPLDPNLRIGGDVGNPIAISEPHSPTAAAI 351


RESULT 9
US-09-738-626-3883
; Sequence 3883, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3883
;   LENGTH: 478
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-3883

  Query Match             9.9%;  Score 164.5;  DB 9;  Length 478;
  Best Local Similarity   28.6%;  Pred. No. 1.2e-07;
  Matches   58;  Conservative   34;  Mismatches   68;  Indels   43;  Gaps   10;

Qy       41 IRSVLQFNRKPELAGETPRIVVITSGKGGVGKTTTTANVGLSLAPYGFSVVAIDADLGLR 100
Db      250 LRTNLQFLN----VGGSSSVFVISSANPGEGKSTTSVNLALALAEAGSRVALIEADLRLP 305

Qy      101 NLDLLLGLENRVNYTCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSKLPMG----FG 156
Db      306 RVSKYLGVEGNAGLT--DILIGKAEVNDVL---QPWGRTQLYYL--PAGPIPPNPSELLG 358

Qy      157 GKALEWLVDALKTRPEGSPDFIIIDCPAGI----------DAGFITAIT------PANE 199
Db      359 SAEMEKVIAEL----EESFDYVIIDAPPALAVTDAAVIGHGKAGILIAVSAGSTKKPELE 414

Qy      200 AVLVTTPDITALRDAD-RVTGLL 221
Db      415 ATLST------LENADANVVGVV 431


RESULT 10
```

```
US-09-815-242-10915
; Sequence 10915, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10915
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-815-242-10915

  Query Match             8.8%;  Score 145.5;  DB 10;  Length 253;
  Best Local Similarity   20.1%;  Pred. No. 3.3e-06;
  Matches   55;  Conservative   51;  Mismatches   72;  Indels   95;  Gaps   12;

Qy      59 RIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDAD--------LGLRN----  --LD 103
           ||: : : ||||||||||  |:|  ||  |  |: || |         :|:          |
Db       3 FIISVANQKGGVGKTTTTVNLGACLANLGKKVLLIDIDAQGNATSGMGVEKPDVAHDVYD 62

Qy     104 LLLGLE---------NRVNY----TCVEVINGDCRLDQALVRDKRWSNFELLCISKPRSK 150
           :|:  |         :| |         :::   :  |   : |: |
Db      63 VLVNEEPITSVVQHTSRENLDIVPATIQLAGAEIELTSMMARESP-------------- 107

Qy     151 LPMGFGGKALEWLVDALKTPPEGSPDFIIIDCPAGIDAGFITAITPANEAVLVTTPDITA 210
                 |:  :| ::       ||::||||  :    | | | |:::|:
Db     108 ---------LKLAIDEVR----DMYDFVLIDCPPSLGHLTINAFT-ASDSILIP------ 147

Qy     211 LRDADRVTGLLECD--GIRDIKMIVNRVR------TDMIKGED-MMSVLDVQEMLGLSLL 261
                  ::|:   :  :  ::| :|         :| |  :::  |  :  ||  ::
Db     148 ----------VQCEYYALEGLSQLLNTIRLVQKHFNPELKIEGVLLTMYDARTNLGAEVV 197

Qy     262 ------------GVIPEDSEVIRSTNRGFPLV 281
                       :|| :  :  : : | |::
Db     198 EEVRKYFPEKVYDTIIPPNVRLSEAPSHGLPII 230


RESULT 11
US-09-815-242-11559
; Sequence 11559, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11559
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Helicobacter pylori
US-09-815-242-11559

  Query Match             8.1%;  Score 134;  DB 10;  Length 264;
  Best Local Similarity   23.3%;  Pred. No. 4.6e-05;
  Matches   63;  Conservative   45;  Mismatches   80;  Indels   82;  Gaps   13;

Qy      60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDAD--------LGLRNLDLLLGLENR 111
           |: : : |||||||||  |:  ||| :   :: || |         || |         ::
Db       5 IIAVANQKGGVGKTTTAVNLAASLAAHEKKILLIDFDPQANATSSLGFR--------RDK 56

Qy     112 VNYTCVEVINGDCRLDQALVR------DKRWSNFELLCISK---------PRSKLPMGFG 156
           ::|    |: |  :: | :::       |   |: |    |           | :|
Db      57 IDYDIYHVLIGRKQISQVILKTQMPFLDLVPSNLGLAGFEKTFYDSQDENKRGEL----- 111

Qy     157 GKALEWLVDALKTRPEGSP----DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALR 212
               | :||     ||      |:|||| |  :    | ::: |:  ::
Db     112 -----MLKNAL-----GSVVKLYDYIIIDSPPALGPLTINSLSAAHSVIIP--------- 152

Qy     213 DADRVTGLLECD  GIRDIKMIVNRVR------TDMIKGEDMMSVLDVQEMLGLSLLGVI 264
                  ::|:   :   |:::| .|         :|   :  . | : | |:  ||:
Db     153 --------IQCEFFALEGTFLLLNTIRMLQKSTNPKLKIRGFLFTMHVPQ LNLT-KGVL 202

Qy     265 PE ----DSEVIRSTNRGFPLVLNKPPTLA 289
           |     |||  | :  |  ::. |   ||
Db     203 AELFKYFDSEFFRDSATGEYIMIPKSVKLA 232


RESULT 12
US-09-767-041-14
; Sequence 14, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
;  APPLICANT:  Smith, Hilda
;  TITLE OF INVENTION:  STREPTOCOCCUS SUIS VACCINES AND DIAGNOSTIC TESTS
;  FILE REFERENCE:  2183-4726
;  CURRENT APPLICATION NUMBER: US/09/767,041
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER:  PCT/NL99/00460
;  PRIOR FILING DATE:  1999-07-19
;  PRIOR APPLICATION NUMBER:  EP98202465.5
;  PRIOR FILING DATE:  1998-07-22
;  PRIOR APPLICATION NUMBER:  EP98202467.1
;  PRIOR FILING DATE:  1998-07-22
;  NUMBER OF SEQ ID NOS:  53
```

```
; SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: Streptococcus suis
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: CPS2C
US-09-767-041-14

  Query Match             7.8%;  Score 128.5;  DB 10;  Length 225;
  Best Local Similarity   23.2%;  Pred. No. 0.00012;
  Matches   56;  Conservative   31;  Mismatches   61;  Indels   93;  Gaps    9;

Qy     52 ELAGETPRIVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDADLGLRN---------L 102
           .|.|    ..| |||  |    ||.|| |.: :: || |:  | :|||  :||        :
Db     29 QLSGADIKVVGITSVKSNEGKSTTAASLAIAYARSGYKTVLVDAD--IRNSVMPGFFKPI 86

Qy    103 DLLLGLENPVNYTCVEVINGDCPLD---------------QALVPDKRWSNFELLCISK 146
             . || . .   |  :.   |  |                 ||:: |  ||| |  :
Db     87 TKITGLTDYLAGT-TDLSQGLCDTDIPNLTVIESGKVSPNPTALLQSK---NFENLLATL 142

Qy    147 PRSKLPMGFGGKALEWLVDALKTRPEGSPDFIIIDCP---AGIDAGFITAITPANEAVLV 203
             |                        |::|:|||      |||  |      |  ||:
Db    143 RRYY-----------------------DYVIVDCPPLGLVIDAAIIAQKCDAMVAVV- 176

Qy    204 TTPDITALRDADPVTGLLECDGIPDIKMIVNPVPTDMIKGEDMMSVLDVQEMLGLSLLGV 263
                        | ::|  :: :|                     :  |  |   |||
Db    177 -------------EAGNVKCSSLKKVK--------------------EQLEQTGTPFLGV 203

Qy    264 I 264
          |
Db    204 I 204

RESULT 13
US-09-738-626-5070
; Sequence 5070, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver  3.0
; SEQ ID NO 5070
;   LENGTH: 261
;   TYPE: PRT
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-5070

  Query Match             7.1%;  Score 118;  DB 9;  Length 261;
  Best Local Similarity   27.8%;  Pred. No. 0.0016;
  Matches   54;  Conservative   31;  Mismatches   67;  Indels   42;  Gaps    9;

Qy     60 IVVITSGKGGVGKTTTTANVGLSLARYGFSVVAIDAD--------LGLRNLDLLLGLENR 111
          |: : : ||||||||:| |:| || | |: :| |         ||:  |
Db      9 IIAMANQKGGVGKTTSTINLGACLAEAGRKVLLVDLDPQGALTAGLGIHYDD-------- 60

Qy    112 VNYTCVE-VINGDCPLDQALVR------DKRWSNFELLCISKPPSKLPMGFGGKALEWLV 164
          |. |  : :.: .  :|||.         |  :| .|  |   :|   |  :  :|
Db     61 VDITVYDLMVDNNSTIDQAIHHTGLPDLDVVPANIDL---SAAEIQLVNEVGRE--QTLA 115

Qy    165 DALKTRPEGSP-DFIIIDCPAGIDAGFITAITPANEAVLVTTPDITALRD ---   -- 213
           ||  ||    |||:.||    .    . |.  |.  ..      .  :||
Db    116 RAL--RPVMKDYDFIILDCQPSLGLLTVNALACAHGVIIPMECEYFSLRGLALLTDTVEK 173

Qy    214 -ADRVTGLLECDGI 226
            |||:   ||  ||
Db    174 VADRLNFDLEILGI 187

RESULT 14
US-09-791-171-94
; Sequence 94, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
;  APPLICANT: ANDERSEN, Peter
;  APPLICANT:  NIELSEN, Rikke
;  APPLICANT:  OETTINGER, Thomas
;  APPLICANT:  RASMUSSEN, Peter Birk
;  APPLICANT:  ROSENKRANDS, Ida
;  APPLICANT:  WELDINGH, Karin
;  APPLICANT:  FLORIO, Walter
;  TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
;  TITLE OF INVENTION:  DERIVED FROM M  TUBERCULOSIS
;  FILE REFERENCE: 670001-2002.1
;  CURRENT APPLICATION NUMBER: US/09/791,171
;  CURRENT FILING DATE:  2001-02-20
;  PRIOR APPLICATION NUMBER: 09/050,739
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 0376/97
;  PRIOR FILING DATE: 1997-04-02
;  PRIOR APPLICATION NUMBER: 1277/97
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/044,624
;  PRIOR FILING DATE: 1997-04-18
;  PRIOR APPLICATION NUMBER: 60/070,488
;  PRIOR FILING DATE: 1998-01-05
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 94
;   LENGTH: 308
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-791-171-94

  Query Match             6.5%;  Score 107.5;  DB 10;  Length 308;
  Best Local Similarity   21.6%;  Pred. No. 0.021;
  Matches   61;  Conservative   46;  Mismatches   97;  Indels   79;  Gaps   14;

Qy     61 VVITSGKGGVGKTTTTANVGLSLARY-GFSVVAIDADLGLRNL------------DLLL 106
           :   ||| |||| ||  :| :||.     ::|:|||  ||                |:|
Db     61 IAVVGLKGGAGKTTLTAALGSTLAQVRADRILALDADPGAGNLADRVGRQSGATIADVLA 120

Qy    107 GLE----NRVN-YTCVEVIN------GDCPLDQALVPDKRWSNFELLCISKPPSKLPMGF 155
           |    |     |  | :|        |     | .  |  |              |
Db    121 EKELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADWH-----FIADPASRF---- 171

Qy    156 GGKALEWLVDALKTRPEGSPDFIIIDCPAG----IDAGFITAITPANEAVLVTTPDITAL 211
                            . .:|| ||     .   :   .    .   |  |   |
Db    172 -------------------YNLVLADCGAGFFDPLTRGVLSTV---SGVVVVASVSIDGA 209

Qy    212 PDADPVTGLLECDGIPDIK----MIVNPVPTDMIKGEDMMSVLDV    QEMLGLSLLGV 263
           |    |   | .|     :::|  :      .||   |  |:        . .     :|
Db    210 QQASVALDWLRNNGYQDLASRACVVINHI----MPGEPNVAVKDLVRHFEQQVQPGRVVV 265

Qy    264 IPEDSEVIRSTNRGF----PLVLNKPPTLAGLA---FEQAAWR 299
           .| |  :   |          |.  |  ||      ||:|  |
```

```
Db    266 MPWDRHIAAGTEISLDLLDPIYKRKVLELAAALSDDFERAGRR 308


RESULT 15
US-09-791-171-70
; Sequence 70, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
;  APPLICANT: ANDERSEN, Peter
;  APPLICANT:  NIELSEN, Rikke
;  APPLICANT:  OETTINGER, Thomas
;  APPLICANT:  RASMUSSEN, Peter Birk
;  APPLICANT:  ROSENKRANDS, Ida
;  APPLICANT:  WELDINGH, Karin
;  APPLICANT:  FLORIO, Walter
;  TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
;  TITLE OF INVENTION:  DERIVED FROM M. TUBERCULOSIS
;  FILE REFERENCE: 670001-2002.1
;  CURRENT APPLICATION NUMBER: US/09/791,171
;  CURRENT FILING DATE:  2001-02 20
;  PRIOR APPLICATION NUMBER: 09/050,739
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 0376/97
;  PRIOR FILING DATE: 1997-04-02
;  PRIOR APPLICATION NUMBER: 1277/97
;  PRIOR FILING DATE: 1997-11-10
;  PRIOR APPLICATION NUMBER: 60/044,624
;  PRIOR FILING DATE: 1997-04-18
;  PRIOR APPLICATION NUMBER: 60/070,488
;  PRIOR FILING DATE: 1998-01-05
;  NUMBER OF SEQ ID NOS: 173
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 70
;   LENGTH: 666
;   TYPE: PRT
;   ORGANISM: Mycobacterium tuberculosis
US-09-791-171-70

  Query Match             6.5%;  Score 107.5;  DB 10;  Length 666;
  Best Local Similarity   21.6%;  Pred  No. 0.06;
  Matches   61;  Conservative   46;  Mismatches   97;  Indels   79;  Gaps   14;

Qy      61 VVITSGKGGVGKTTTTANVGLSLARY-GFSVVAIDADLGLRNL------------DLLL 106
           : :   ||| |||| || :| :||:     ::|:||| |  ||             |:|
Db     419 IAVVGLKGGAGKTTLTAALGSTLAQVRADRILALDADPGAGNLADRVGRQSGATIADVLA 478

Qy     107 GLE----NRVN-YTCVEVIN------GDCRLDQALVRDKRWSNFELLCISKPRSKLPMGF 155
             |    | :  :| |  :|       :    |  : |  |       |: | |:
Db     479 EKELSHYNDIRAHTSVNAVNLEVLPAPEYSSAQRALSDADWH-----FIADPASRF---- 529

Qy     156 GGKALEWLVDALKTRPEGSPDFIIIDCPAG----IDAGFITAITPANEAVLVTTPDITAL 211
                          : :: || ||    :  | :: :   :  |:| :  |
Db     530 -------------------YNLVLADCGAGFFDPLTRGVLSTV---SGVVVVASVSIDGA 567

Qy     212 RDADRVTGLLECDGIRDIK----MIVNRVRTDMIKGEDMMSVLDV----QEMLGLSLLGV 263
           : |       |  ·| :|:     :::| :    : ||  ::| |:    :: :    : |
Db     568 QQASVALDWLRNNGYQDLASRACVVINHI----MPGEPNVAVKDLVRHFEQQVQPGRVVV 623

Qy     264 IPEDSEVIRSTNRGF----PLVLNKPPTLAGLA---FEQAAWR 299
           :| |  :   |        |:   |  ||      ||:|  |
Db     624 MPWDRHIAAGTEISLDLLDPIYKRKVLELAAALSDDFERAGRR 666


Search completed: April 16, 2003, 09:09:25
Job time : 22 secs
```